NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

NEGATIVE BIAS TEMPERATURE INSTABILITY (NBTI) EXPERIMENT

by

Christopher Mark Schuster

June 2006

Thesis Advisor:	Todd R. Weatherford
Second Reader:	Andrew A. Parker

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2006	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE: Negative Bias Temperature Instability (NBTI) Experiment			**5. FUNDING NUMBERS**
6. AUTHOR(S) Christopher Mark Schuster			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

The phenomenon known as Negative Bias Temperature Instability (NBTI) impacts the operational characteristics of Complementary Metal Oxide Semiconductor (CMOS) devices, and tends to have a stronger effect on p-channel devices. This instability is observed with an applied "on" biasing during normal operation and can be accelerated with thermal stress. A normal applied electrical bias on CMOS transistors can lead to the generation of interface states at the junction of the gate oxide and the transistor channel. The hydrogen that normally passivates the interface states can diffuse away from the interface. As a result, the threshold voltage and transconductance will change. These interface states can be measured to determine the susceptibility to NBTI of the devices. For this purpose, a charge pumping experiment and other On-the-Fly techniques at certain temperatures can provide the interface state density and other valuable data. NBTI can impact current technological fabrication processes, such as those provided to the government from IBM. This paper explains this testing of current submicron transistor technology that will be used for military applications.

14. SUBJECT TERMS NBTI, On-the-Fly, Semiconductor Reliability, IBM Trusted Foundry Program, 130nm, Military Electronics			**15. NUMBER OF PAGES** 75
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

NEGATIVE BIAS TEMPERATURE INSTABILITY (NBTI) EXPERIMENT

Christopher Mark Schuster
Ensign, United States Navy
B.S., United States Naval Academy, 2005

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 2006

Author: Christopher Mark Schuster

Approved by: Todd R. Weatherford
Thesis Advisor

Andrew A. Parker
Second Reader

Jeffery B. Knorr
Chairman, Department of Electrical and Computer Engineering

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The phenomenon known as Negative Bias Temperature Instability (NBTI) impacts the operational characteristics of Complementary Metal Oxide Semiconductor (CMOS) devices, and tends to have a stronger effect on p-channel devices. This instability is observed with an applied “on” biasing during normal operation and can be accelerated with thermal stress. A normal applied electrical bias on CMOS transistors can lead to the generation of interface states at the junction of the gate oxide and the transistor channel. The hydrogen that normally passivates the interface states can diffuse away from the interface. As a result, the threshold voltage and transconductance will change. These interface states can be measured to determine the susceptibility to NBTI of the devices. For this purpose, a charge pumping experiment and other On-the-Fly techniques at certain temperatures can provide the interface state density and other valuable data. NBTI can impact current technological fabrication processes, such as those provided to the government from IBM. This paper explains this testing of current submicron transistor technology that will be used for military applications.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

The author would like to thank Todd Weatherford for his experience and guidance during this work. As well, I would like to recognize Gamani Karunasiri, for his support made this project possible.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

NBTI is a process-dependant phenomenon mostly affecting PMOS devices that occurs as a result of a number of factors. A lattice mismatch between the bulk Si and the SiO_2 gate leads to the creation of dangling bonds. This effect creates non-ideal effects in the transistor that are called interface states. These interface states can trap charge and change the operational characteristics of the device. These dangling bonds can be terminated with Hydrogen due to the use of steam to grow the gate oxide, which has the effect of passivating them and rendering them inert. However, under an electrical field and the influence of a thermal input, the hydrogen can disassociate from the bonds and diffuse through the material. As a result, the device will change its operating parameters.

This instability is affected by the magnitude of the applied voltage stress and can be accelerated with temperature. Due to the charge accumulation, the PMOS device experiences a reduction in current as the threshold voltage to turn the device "on" shifts. After long periods of the stress, the device will eventually turn "off" and become inoperable. This can lead to timing issues in circuits as well as catastrophic failures.

NBTI is a problem for the military for a number of reasons. First, the military uses commercial suppliers for its microelectronic needs, such as from the IBM Trusted Foundry Program. Current commercial technology is affected by this instability. Second, the military operates in harsh environments which may exacerbate the NBTI effect. As a result, this research is needed to determine the level of susceptibility of the IBM fabrication process that will be used for military applications.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. OBJECTIVE

The goal of this research is to provide magnifying information about NBTI and its effects on the IBM Trusted Foundry 130nm process for military applications. This project applies the On-the-Fly measurement technique [1-3] to a p-channel Metal-Oxide-Semiconductor (PMOS) transistor on an Air Force Research Laboratories (AFRL) test structure to extract pertinent information on device degradation. It includes establishing test procedures that operate integrated temperature monitoring and control. If the process is susceptible to NBTI, such information can be critical to designers and military contractors should they choose to take advantage of the IBM CMOS process.

B. BACKGROUND

This thesis is part of a project involving David Alexander from AFRL in Albuquerque, New Mexico at AFRL Kirtland Air Force Base. This experiment makes use of a test structure designed by Don Pierce of Sandia Technologies and manufactured by Sandia National Laboratories for the specific purpose of reliability testing of integrated circuits. The design also includes test structures to conduct experiments for hot carrier injection, electro-migration, plasma damage, and gate oxide integrity in addition to the NBTI structures. Additional work on the NBTI evaluation is also under way by LT Isahi Cortes, USN. This research addresses the reliability of a current "trusted" semiconductor fabrication facility that provides for the manufacture of devices for Department of Defense (DoD) consumption.

The motivating factors behind this research can fall into the following three categories, all of which hold special interest to the military. First, there is the fact that the DoD has specific and sometimes stringent reliability needs for its microelectronic components. Second, the military is not as large a player in the semiconductor market as it once was, and as a result may not be able to fill its reliability needs in a cost effective manner. The third important driving factor is that the government is aware of the aforementioned discrepancies and is taking steps to remedy this situation. These areas will be discussed in the following sections.

1. DoD Reliability Needs

The United States' security forces have specific needs and requirements for electronic equipment, as they operate in adverse climates and under sensitive security conditions. Mission requirements can place them in locations that range from the undersea environments of submarines to the furthest heights of geostationary orbit aboard satellites and back to earth in the computers and surveillance equipment of the National Security Agency. On top of these possibly adverse climes, the DoD and other interested agencies look to have equipment and electronics that will withstand the normal burdens of operation providing consistent support to the user for the duration of the product's time in use. Though a system or product may undergo upgrades, some lifespans can be on the order of decades. For example, the B-52 Stratofortress aircraft is scheduled to be in operation for approximately 84 years through service extensions [4].

a. Military Standards and Performance Ratings

In the past, these requirements were met via the use of military standards and qualified parts, with an example of MIL-STD-883E for microcircuitry. This standard provided information to determine if a part is "…suitable for use within Military and Aerospace electronic systems including basic environmental tests to determine resistance to deleterious effects of natural elements and conditions surrounding military and space operations…," [5] ensuring usability for the Department of Defense. Other standards and performance specifications outline desired characteristics for components to be used in DoD systems. However, the availability and applicability of parts that meet these standards and performance specifications has been greatly affected by the changes in the current semiconductor market. As consumer demand took priority over DoD needs, the government had to adjust its requirements to be able to find vendors that were cost effective and available.

b. Secure Suppliers

Another category of reliability that can pose a problem to current and future military programs is the one not on an actual device operational and design characteristics per se, but that of a reliable support and supply structure for military electronics. If the DoD cannot acquire necessary parts due to reliance on international suppliers, this can be a liability for operational security and readiness. Moreover, if a

secure “trusted” source of microelectronics cannot be found, yet again this is an issue for the military. More background is provided in the section on the semiconductor market, as this a definite concern of the DoD.

2. DoD and the Semiconductor Market

Although the military has specific needs for electronics in the realm of reliability and security, it only commands a small fraction of the overall demand. The Defense Science Board estimates DoD consumption on the order of 1-2% of global supply [6]. Not only does this mean that the DoD has less of a say in the direction of development of technology, but also that it will be subject to the changes in market trends. The following chart illustrates a breakdown of consumption for the different market areas.

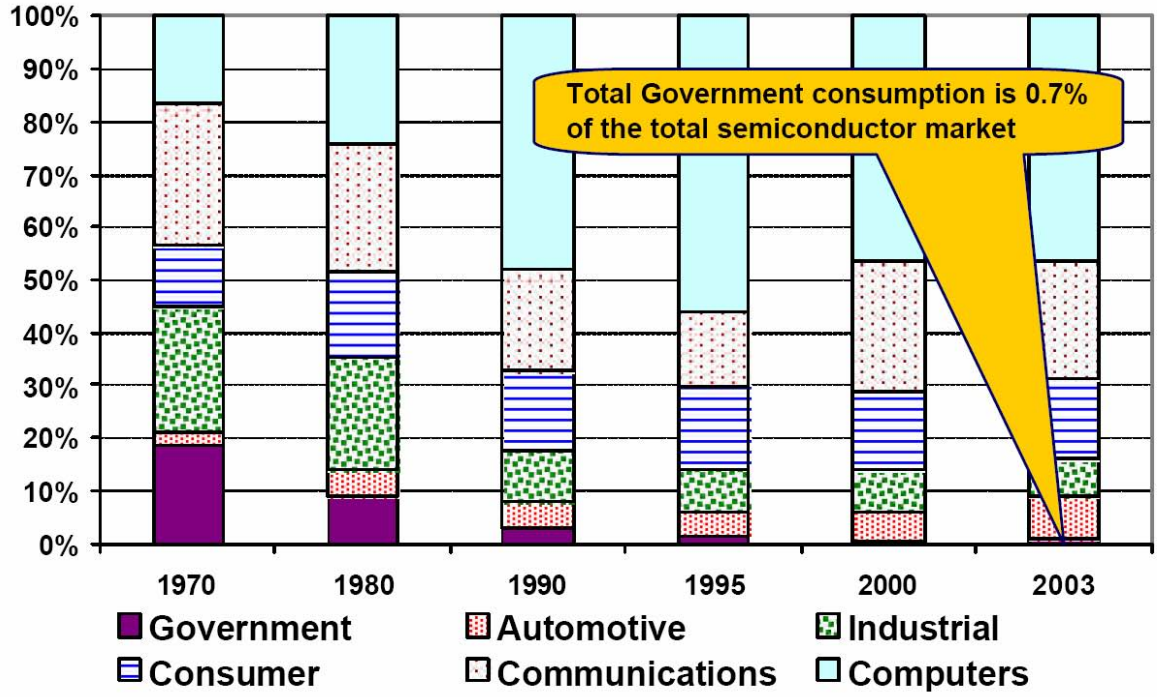


Figure 1. Semiconductor Consumption [7].

This makes acquisitions for the military more difficult for certain specific needs because of the reduced leverage, which fall in line with what should be primary concerns of the government. These concerns are that of device reliability and that of the reliability of the companies manufacturing the devices.

Extended device reliability is not necessarily a top priority for microchip manufacturers. They are concerned with such problems as infant mortality rates and

yields so that the customer can get a functioning product. However, this can be independent of the overall lifetime of the device. By virtue of the manufacturing techniques the devices can be susceptible to damage from electromigration, electrostatic discharge, hot carrier injection, tin whiskers in lead-free solder and other reliability issues. This can reduce chip reliability to such lengths of time of two years and even less in some instances.

A separate issue that the government is facing is that of the reliability of supply. As more companies are opting to send their manufacturing to off-shore locations to take advantage of cheaper operational costs [6], fewer on-shore locations remain. This can manifest itself as a reliability concern in a number of ways. Primarily, if the government wishes to manufacture circuitry that has specific national security concerns associated with it, such as guidance systems for missiles or processors for NSA communication equipment, it must send that information on how to build that equipment to a foreign nation. This raises the chance of getting "compromised microelectronic circuits" [6] which could be either referring to reverse engineering of the sensitive technology or the introduction of a "Trojan horse" into the returned manufactured system. In a different fashion, it is possible that a supply of necessary components could be interrupted during war or peacetime due to political or other concerns.

3. Government and DoD Reaction

In reaction to the threat of losing opportune on-shore sites to produce trusted circuitry, the government has begun to take some small steps to provide for electronics that can be used for national security and military applications. There are multiple programs underway to explore the possible options available for a remedy. Some sources suggest the formation of a Consolidated DoD Semiconductor Foundry [7, 8], yet admit of its costliness and the possible negative impact on the national microelectronics industry [8]. The Defense Science Board published the most recent report in 2005 investigating this problem and proposes possible solutions. This report calls for a long term strategy to deal with the supply and reliability problems for today and into the future. However, there exists a short-term solution that relies upon continental United States manufacturers.

a. The Defense Trusted Integrated Circuits Strategy (DTICS)

In October of 2003, Deputy Secretary Wolfowitz provided guidance for the DTICS by providing directional goals [6]. These goals lead to the development of the Trusted Affairs Programs Office (TAPO) and the beginning of an industrial base for secure production of trusted equipment. It is in accordance with this strategy that the TAPO established contractual relationship with International Business Machines (IBM) for "trusted" microelectronics [9].

During the fiscal year of 2004, a small portion of the defense budget was appropriated to pay for use of semiconductor wafer fabrication facilities at IBM in Vermont. Though the drive for the funding primarily was a result of the NSA, the defense budget allocation allows for use by all DoD parties. These facilities would provide for manufacturing of current technologies, using the standard technologies of Bulk CMOS, RF CMOS, and Silicon Germanium (SiGe) BiCMOS that IBM provides [9]. The following figure shows a breakdown of the features these technologies provide.

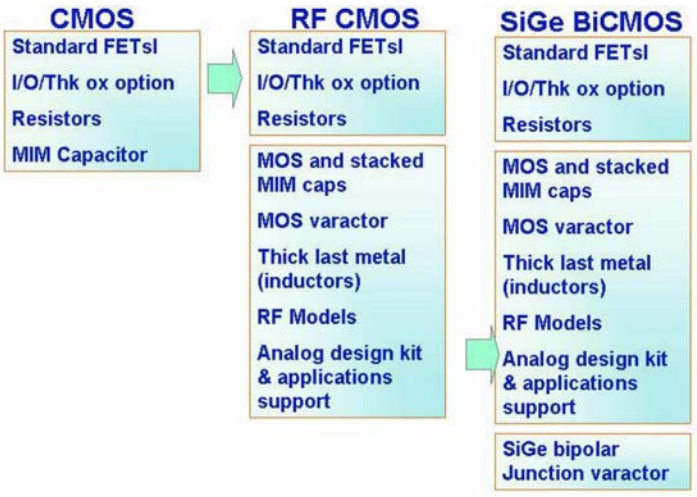


Figure 2. IBM Technology Offerings [10]**.**

Though the agreement will provide a method for manufacturing advance microelectronic circuitry it is specifically only guaranteed to meet the reliability standards of "IBM's standard commercial processes, and as such may not meet military environmental or radiation requirements," [9]. With this fabrication facility open to the

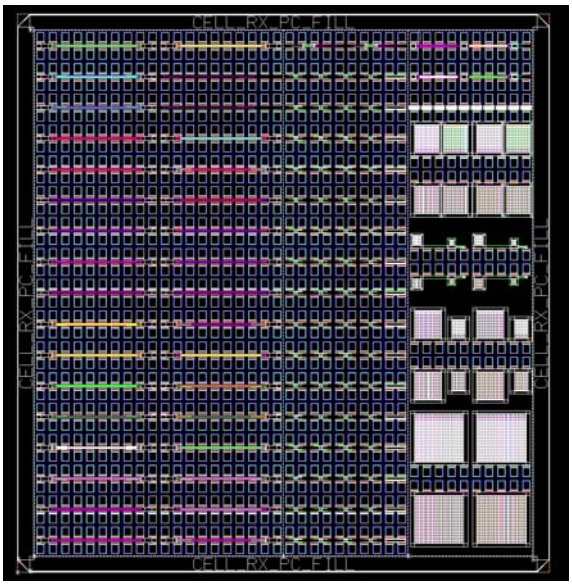
use of the government, the problem of military supply is in the initial stages of repair. However, the reliability of the devices themselves is still of great import, and as such, must be known.

II. AFRL TEST STRUCTURE

The AFRL, in association with Sandia Technologies, produced the reliability test structure that is used in this work. It is a comprehensive design that incorporates experiments to explore electromigration, hot carrier effects, NBTI, and other test structures. It is built using the IBM process CMRF8SF, which is a 130nm gate length CMOS process.

A. DEVICE OVERVIEW

The die is a square that is approximately 5x5 mm. It is a bare die with bond pads that is ready for testing via probing with tips or a probe card. Though they are provided unbound, they can fit into a 24-pin DIP package or other carrier and can be wire bound. The chip layout is shown in the figure below. The NBTI structure is located within the top two rows on this drawing layout in Figure 3. A depiction of which area of the die is allocated for each experiment is included on the right for clarification [11]. The prefix of EM indicates an electro-migration test, HC is for a hot carrier experiment, CAP designates an oxide integrity test, and ILD specifies an interlayer dielectric test.



EM_M3_p2u	EM_M4_p2u	NBTI	ILD_M4_M5_M6_COMB
EM_M8_3u	EM_PCONT	HC_D_P_HV_p24	ILD_M1_M2_M3_COMB
EM_M8_p64u	EM_CONT_N_P	HC_D_N_HV_p24	CAP_TRANS
EM_M7_3u	EM_V7_U	HC_C_P_HV_p24	
EM_M7_p64u	EM_V7_L	HC_C_N_HV_p24	CAP_FIA
EM_M6_3u	EM_V6_U	HC_B_P_HV_p24	
EM_M6_p32u	EM_V6_L	HC_B_N_HV_p24	
EM_M5_3u	EM_V5_U	HC_A_P_HV_p24	CAP_A_SMALL
EM_M5_p32u	EM_V5_L	HC_A_N_HV_p24	
EM_M4_3u	EM_V4_U	HC_D_P_LV_p12	
EM_M4_p32u	EM_V4_L	HC_D_N_LV_p12	CAP_A_MIDDLE
EM_M3_3u	EM_V3_U	HC_C_P_LV_p12	
EM_M3_p32u	EM_V3_L	HC_C_N_LV_p12	
EM_M2_3u	EM_V2_U	HC_B_P_LV_p12	
EM_M2_p32u	EM_V2_L	HC_B_N_LV_p12	CAP_A_BIG
EM_M1_3u	EM_V1_U	HC_A_P_LV_p12	
EM_M1_p24u	EM_V1_L	HC_A_N_LV_p12	

Figure 3. AFRL Test Structure.

B. NBTI STRUCTURE SPECIFICS

In the NBTI structure, there are twenty bond pads available for connection to the circuitry in the device. The design incorporates two p-channel devices, two polysilicon heaters, and two thermistors. Each PMOS device has its own thermistors and heater built directly next to the device. One of the MOSFET devices has a protection diode integrated into the design. The thermistors are designed to allow for a Kelvin connection in order to obtain a more accurate measurement of the resistance. As a result, each thermistor has four pads associated with it instead of two.

1. PMOS Device Details

The transistors themselves are built using 0.13µm technology. The design length of the gate is 0.12µm and the width is 10µm [11]. They are low voltage devices that operate under a supply voltage of 1.5V with gate voltages in the neighborhood of $-2.2V$. The following chart describes the operating characteristics for the IBM process at the 130nm node. The chart reveals that the devices have an operating threshold voltage of $-0.3V$ for the PMOS devices.

CMOS specifications (common to 130-nm technology platform)	
Lithography	130 nm
Voltage (V_{DD})	1.2 V or 1.5 V
Additional power supply options	2.5 V / 3.3 V I/O
Standard NFET / PFET	
L_{min}	0.12 µm
L_p	0.09 µm
V_{tsat}	0.355 V / -0.300 V
I_{Dsat}	530 µA/µm / 210 µA/µm
I_{off}	300 pA/µm / 350 pA/µm
T_{ox}	2.2 nm

Note: Specifications given for 1.2 V (nominal) at 25°C.

Figure 4. IBM 130nm Node Data [10]**.**

2. Thermal Control System Details

For each transistor, the heater and thermistor are located physically near the devices for accurate thermal measurements. The polysilicon heaters have a resistance of 20Ω. The thermistor itself has a resistance near 20Ω at 25˚C. This value of resistance is extracted from the combination of the force and sense line resistances. Measured from

the sense line, the resistance will be on the order of 60Ω at 25˚C. From the force line, the resistance is 40Ω at 25˚C. The following figure shows a simplified circuit schematic of the NBTI test structure.

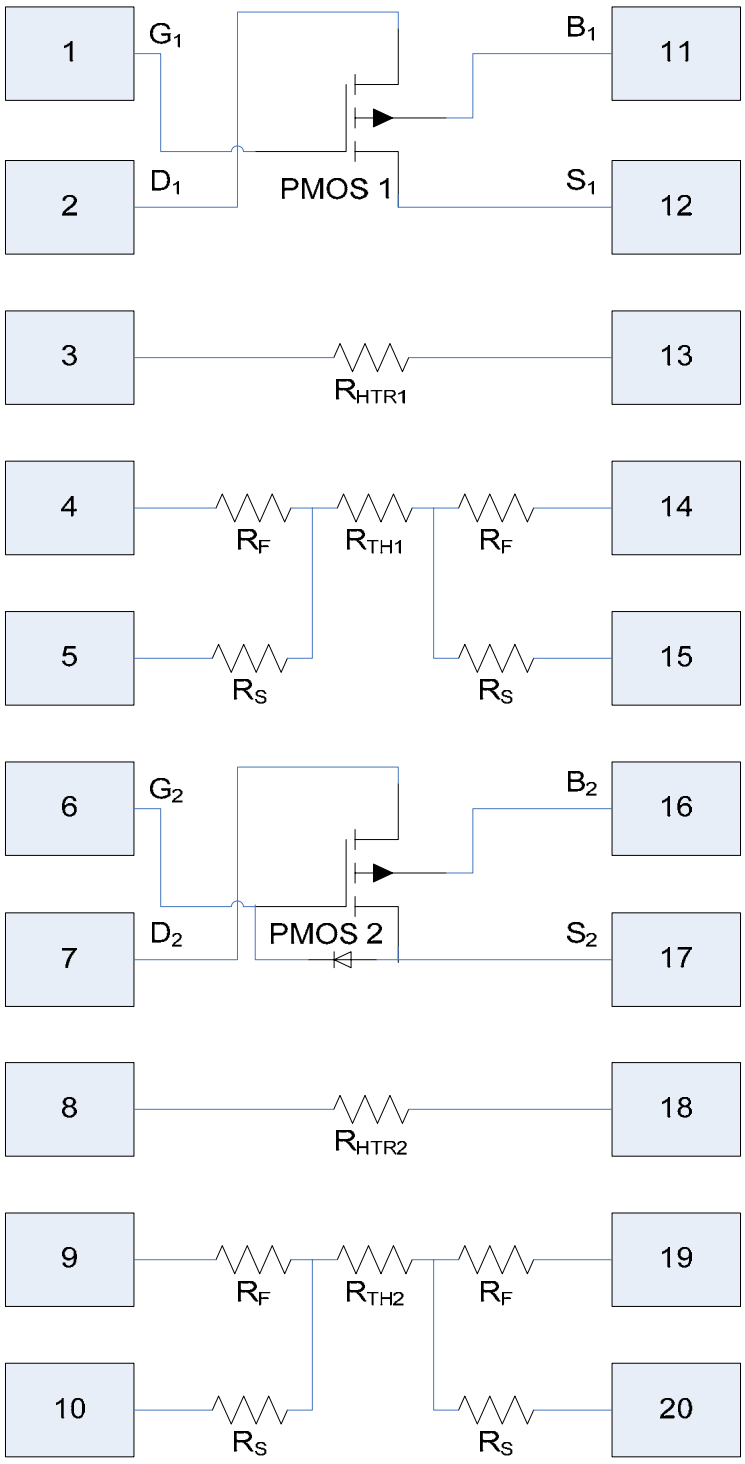


Figure 5. NBTI Test Structure Schematic.

THIS PAGE INTENTIONALLY LEFT BLANK

III. NEGATIVE BIAS TEMPERATURE INSTABILITY (NBTI)

A. INSTABILITY INTRODUCTION

NBTI causes a shift in the threshold voltage. This is occurs when a bias is applied to put the PMOS device channel into inversion. It also can be aggravated when a device is subjected to an increased temperature and excessive voltages. It is a nonlinear instability that arises from charge interactions at the interface of the silicon body and the gate oxide. There are multiple factors that contribute to effects of this instability. They are related to the size of the device, specifically to that of the physical features, the process used to produce the oxide on the gate due to the hydrogen passivation of interface states, and that of the device type. NMOS transistors are not significantly affected. To fully understand the impact of the instability, it is important know what aspects of operation can be affected. To this end, this section discusses the normal operation of a PMOS.

1. Ideal PMOS Operation

A generic cross-section of a p-channel device is shown below in Figure 6 for an enhancement type PMOS with four external connections, as well as the conceptual circuit component. These will be used in the following discussion of DC operation for an enhancement mode p-channel device.

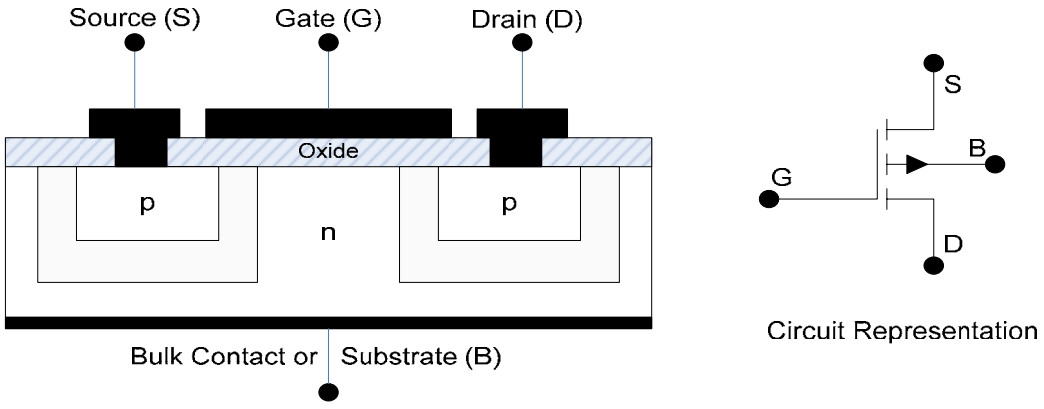


Figure 6. PMOS Cross-section and Circuit Component [12].

The first portion of this section addresses the important and basic characteristics of a PMOS device at the level of operation. It provides background that can easily be

applied to simple experimentation. The latter part addresses the operation of the devices from an in-depth perspective to investigate the areas where the degradation can play a role.

a. Basic Operation

The PMOS and all MOSFETs in general work as a switch. In most operational states, a differential voltage is applied across the "Source" and "Drain" connections. When the proper voltage is applied to the gate, a channel of charge carriers will form under the gate and current will flow. The voltage above which current will flow is referred to as the threshold voltage V_T. It is measured in potential relative to the source potential [13]. The required gate voltage to cause current flow is related to the physical characteristics of the device. Figure 7 shows an experimental setup to determine the threshold voltage. This graph shows the threshold voltage as taken from the extrapolation of the linear portion of the drain current. Another method uses the same technique but plots the data as the square root of the drain current.

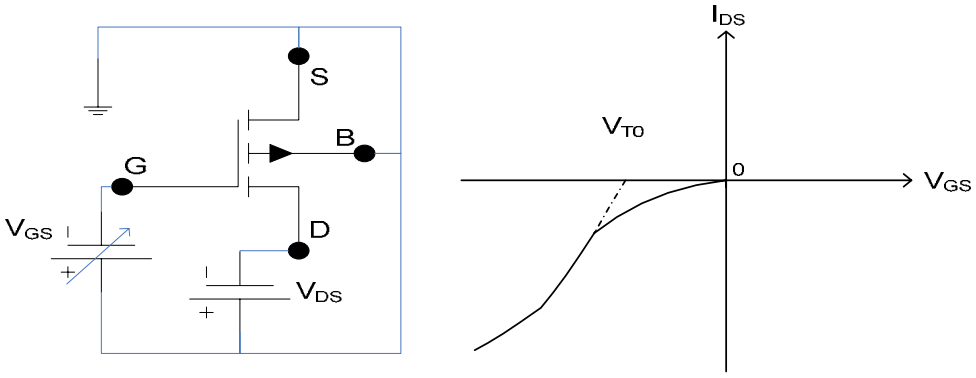


Figure 7. Threshold Voltage Extraction Bias and Curve [14]**.**

Different biases can yield different data results that can provide information about the physical characteristics of the device. A commonly used plot of the drain current and the drain bias can show the mode of operation. Depending on the combination of biasing the device can operate in the cutoff region, the triode region, or the saturation region [13]. These different regions establish where the device is "on" and "off". This plot is generated by sweeping the drain bias at different and constant gate voltages. Figure 8 shows a combination of the biases and the resulting operating regions.

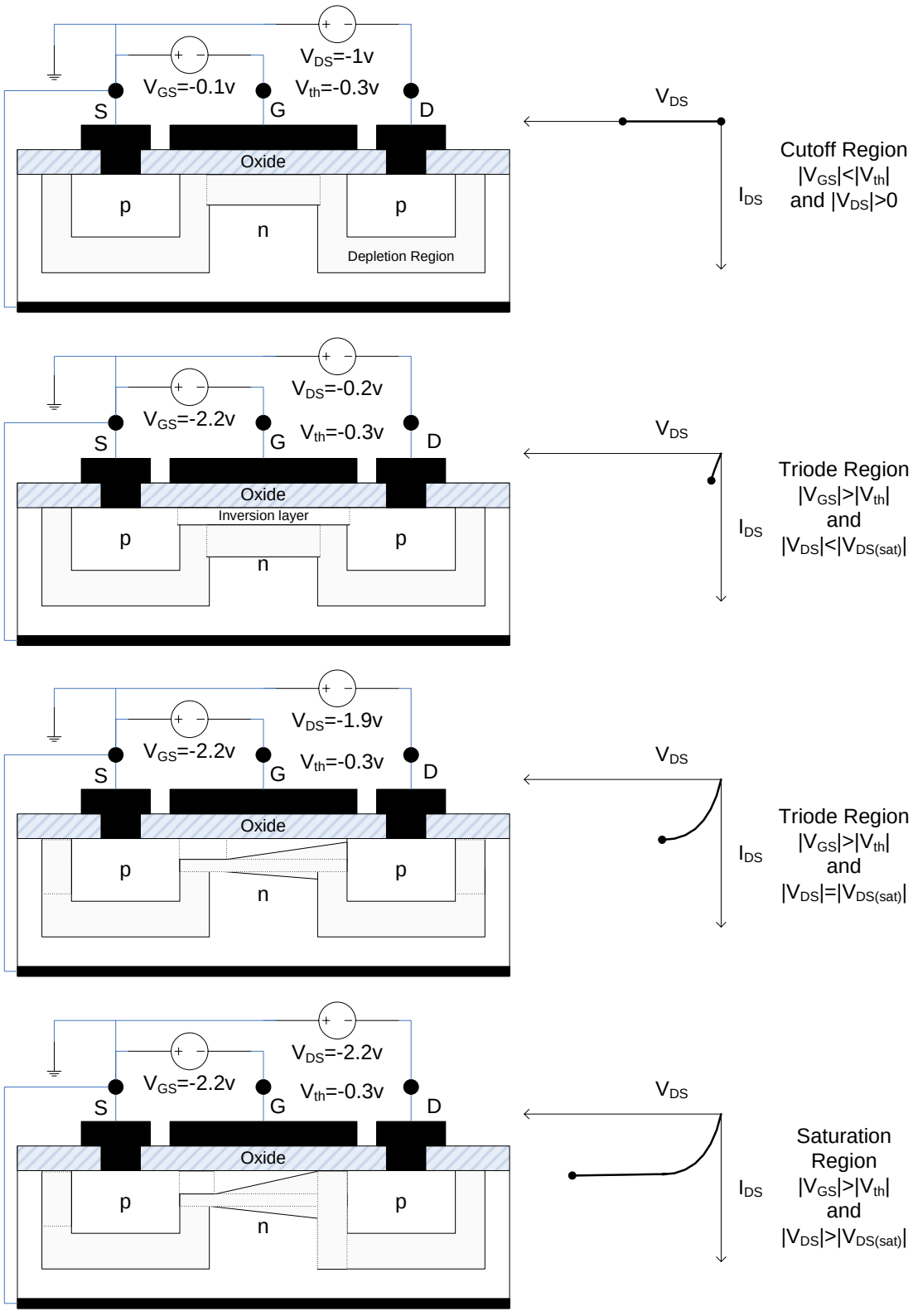


Figure 8. PMOS I_D versus V_{DS} Biases [12].

The cutoff region is the point where the applied gate voltage does not provide enough potential to repel the negative charge carriers in the channel to allow current flow. Thus the gate voltage is less than the threshold voltage and no significant amount of current will flow through the channel. Due to the fact that a voltage is applied, a depletion layer forms in the channel. As a negative voltage is applied to the gate, the majority carrier of the n-doped channel is repelled deeper into the bulk n material.

Once the voltage applied to the gate exceeds the threshold voltage, enough of the majority carriers in the n-type material will be repelled to form an inversion layer. This is the resultant p-channel that provides the capacity to carry current. At this point, the transistor is "on" and can conduct. As a small voltage is applied across the source and drain, the current will increase in a linear fashion with the applied voltage. However, as the applied negative voltage to the source and drain increases in magnitude, a channel effect begins to saturate the current. This controls the resistance of the channel.

Different locations along the length of the channel will be at different voltages as the applied voltage is dropped across the channel. The channel voltage does not affect the threshold voltage, but due to the additional bias in the channel, the actual applied differential voltage is reduced near the drain. When the difference between the gate voltage and the drain voltage is equal to the threshold voltage, the channel is "pinched off" at the drain. The second and third graphs in Figure 8 depict operation in this fashion, and it is called the triode region. Past the pinch-off point in the channel, the current is saturated and the current only increases slightly with an increasingly negative drain voltage. This region is called the saturation region and is depicted in the fourth group in Figure 8.

b. Energy Band Diagram

The electron band energy diagram can provide insight into the operation of the device. These diagrams provide a visual representation of the electron energies in the gate and the channel. It is taken as a cross-section of the device starting at the gate on the leftmost part of the diagram. As the 'x' distance increases, the energies are drawn for the gate oxide, the channel, and the bulk material. The following figure shows the ideal characteristic for a PMOS transistor in the flatband, depletion, and onset of inversion conditions [15]. This reflects the relative dopant concentrations and other features of the

device. The horizontal axis is the distance of depth into the transistor, and is a common depiction of MOS devices. In the inert flat-band state, the Fermi level of the metal is on the same level as the Fermi level of the channel. Specific to this device and showing that it is a PMOS, the bulk material is n-type silicon.

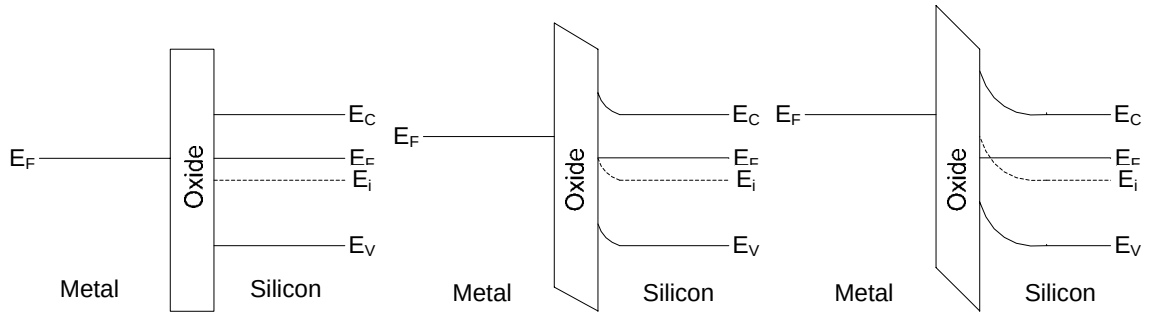


Figure 9. Ideal Band Diagrams [15]**.**

When a voltage is applied, the desired effect is an inversion of the majority carrier in the channel space adjacent to the gate oxide as is visually depicted in Figure 8. To invert this region, the electrons must be repulsed, revealing the ion with a positive charge. Thus the applied voltage must be more negative to repel the like carriers. The point where the transistor will begin to conduct occurs when the threshold voltage is reached. There are a few contributing factors that can determine this voltage. The potentials of the devices at different depths affect the threshold point in the device. Figure 10 shows a representation of potential at the surface of the interface and the bulk material.

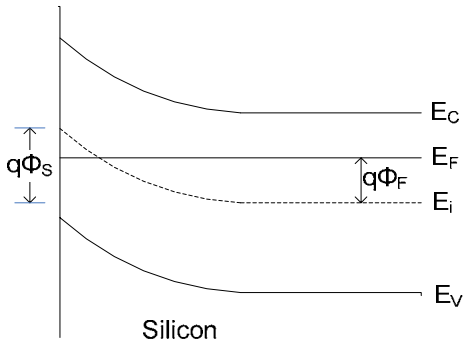


Figure 10. Potential Diagram from the Oxide to the Bulk [15]**.**

When the voltage is applied to the gate, it will be partly dropped across the oxide and the rest across the semiconductor. The magnitude of the voltage drop across the oxide relates to its thickness as well as the dielectric properties of the two materials. The voltage drop across the semiconductor relates to the change in potential from the surface on into the bulk. The following equation shows the voltage relationship for this situation [15].

$$V_G = \Delta\Phi_{Oxide} + \Delta\Phi_{Semiconductor} \tag{3.1}$$

The V_G term is the voltage applied to the gate. The $\Delta\Phi$ terms represent the change in potential as dropped across the oxide and semiconductor respectively. Since the oxide drop is related to the properties discussed above, it can be written as the following equation. This assumes that the oxide is ideal [15].

$$\Delta\Phi_{Oxide} = \frac{K_S}{K_O} x_0 \varepsilon_S \tag{3.2}$$

The terms K_S and K_O relate to the dielectric properties where the 'S' represents that of the semiconductor and the 'O' refers to the oxide. The ε_S term defines the electric field at the interface, and x_0 is the thickness of the gate oxide [15].

Two important potential values are that of the Fermi potential of the bulk material, which is shown at the far right of Figure 10 as the Φ_f term, and that of the oxide surface potential depicted as Φ_S. They are both dependant upon the doping levels, where the bulk level is related to the base doping of the substrate material and surface level relates to the doping caused by the gate voltage. The Φ_S term is a reduction of Φ_{Semi} where the bulk is ground. Taken together, the gate voltage is related to the other potentials by Equation 3.3 [15].

$$V_G = \frac{K_S}{K_O} x_0 \varepsilon_S + \Phi_S \tag{3.3}$$

Depending on the relationship of the electric field at the surface, which can be given in the following equation [15], the full version of the potential relationship with the gate voltage is below [15].

$$\varepsilon_S = \sqrt{\frac{2qN_A}{K_S\varepsilon_0}\Phi_S} \tag{3.4}$$

$$V_G = \frac{K_S}{K_O}x_0\sqrt{\frac{2qN_A}{K_S\varepsilon_0}\Phi_S} + \Phi_S \tag{3.5}$$

The electric field is dependant upon the charge q, the concentration of dopants N_A, the permittivity of free space ε_0, and the other previously defined variables. The device is said to turn on after the onset of inversion, where the surface potential is twice that of the bulk material [15]. This yields the final equation for the threshold voltage for the ideal p-channel case.

$$V_T = 2\Phi_F - \frac{K_S}{K_O}x_o\sqrt{\frac{4qN_D}{K_S\varepsilon_o}\left(-\Phi_F\right)} \tag{3.6}$$

Equation 3.6 takes into account the fact that the dopant will act as a donor for a p-channel, n-bulk device, and therefore the contribution to the electric field from dopants is taken as the number of donors N_D. From this equation of the ideal case and an understanding of the band diagrams, it is possible to visualize the operation of the device. It is also not too difficult to infer how a device will act in a non ideal situation.

B. DEVIATION FROM IDEAL OPERATION

Since this event occurs as a result of charge accumulation at the interface it is relatively easy to visualize the steps of the instability. This section addresses the non-ideal operation of a transistor for comparison to the ideal effects. There are multiple ways in which the non-ideal MOSFET can differ from the ideal case. Defects that affect the threshold voltage are important to consider. One issue is the difference in the workfunctions of the metal contact and the semiconductor. Other problems arise from defects that add charge in the oxide in some form. There are four types of charge events

that occur, mobile ion charges, oxide trapped charge, interface trapped charge, and fixed charge [15, 16]. The interface charges receive treatment in the following sections as they relate to NBTI.

1. Work Function Difference

In general, the work function of a material is measured from the Fermi energy level to that of the vacuum energy of a free electron [12]. This is where the electron would be free of the influence of the material. This required energy needed to remove the electron from a material to a free state differs from material to material. When two materials that have a different work function are brought into electrical contact, the higher energy electrons will flow to the lower potential areas to balance the Fermi energy between the two materials. If there is a material that impedes the flow of electrons, such as an insulating oxide layer, the difference in potential will create a field to accommodate the adjustment in the Fermi levels. In the ideal case for a MOSFET, the workfunctions would be equal and would be described by the left portion of Figure 11. Due to the real differences, the actual flat-band state would appear in the form of the second half of Figure 11.

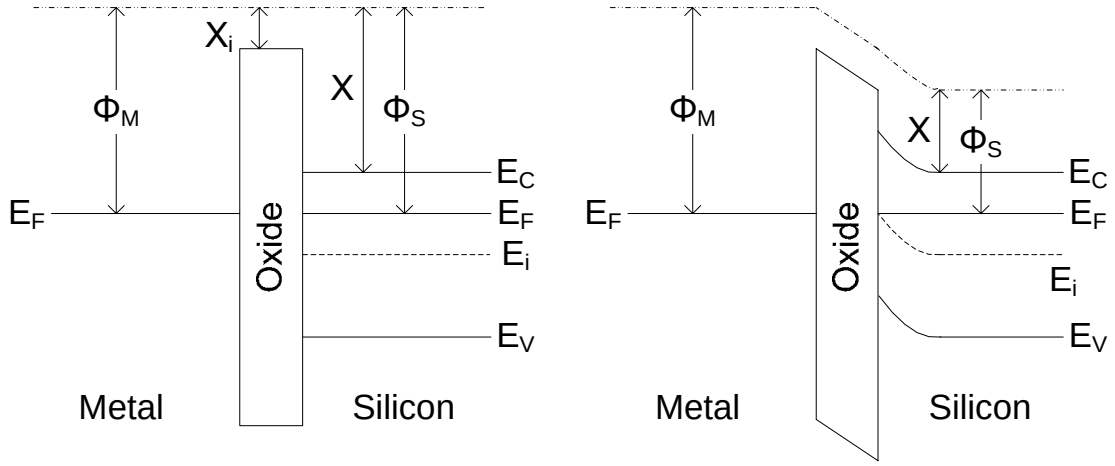


Figure 11. Effect of Workfunction Difference.

This difference leads to a shift in the relationship of the gate voltage to the Fermi level and surface potential at the oxide. The difference is defined in the following equation [15].

$$\Phi_{MS} = \frac{1}{q}\left(\Phi_M - \Phi_S\right) \tag{3.7}$$

When this is added to the ideal case, it leads to the non-ideal case that incorporates the workfunction differences as shown below.

$$V_G = \frac{K_S}{K_O} x_0 \sqrt{\frac{2qN_A}{K_S \varepsilon_0} \Phi_S} + \Phi_S + \Phi_{MS} \tag{3.8}$$

Though this difference will cause a shift from normal operation, it is a very predictable non-linearity. It will not change as the device is operated as it is a specific property related to the materials themselves. The workfunctions will not change with different biases. As such, it is taken into account.

2. Addition of Oxide Charge

The charges trapped in and around the oxide are of an increased concern. There are there are two specific types that are necessary to understand that relate to NBTI. These are the fixed charges and the interface charges. These tend to be process dependant and can be malleable quantities.

a. Fixed Charge

The fixed charge is charge that is trapped near the interface, but not directly adjacent to it. The resultant shift from this trapped charge is observed to be negative, and therefore the trapped charge itself must be positive. This addition to the non-ideal nature stems from the presence of excess ionic silicon in the vicinity of the interface [15]. The contribution is shown in the following equation.

$$\Delta V_G = -\frac{Q_F}{C_o} \tag{3.9}$$

The C_o term is defined as the capacitance of the gate oxide, and the Q_F term is the fixed oxide charge per unit area of the gate [15].

b. Interface Charge

The generation of interface charges is an artifact of the oxidization process. The oxygen provided via steam bonds with the silicon as it is exposed to the crystal, but due to differences in lattice constants and bonding orientations, there will be unpaired electrons. These states are referred to as P_b centers [17, 18]. Different crystal

orientations give rise to different types of interfacial traps, and a graphical representation of possible orientations is shown in Figure 12.

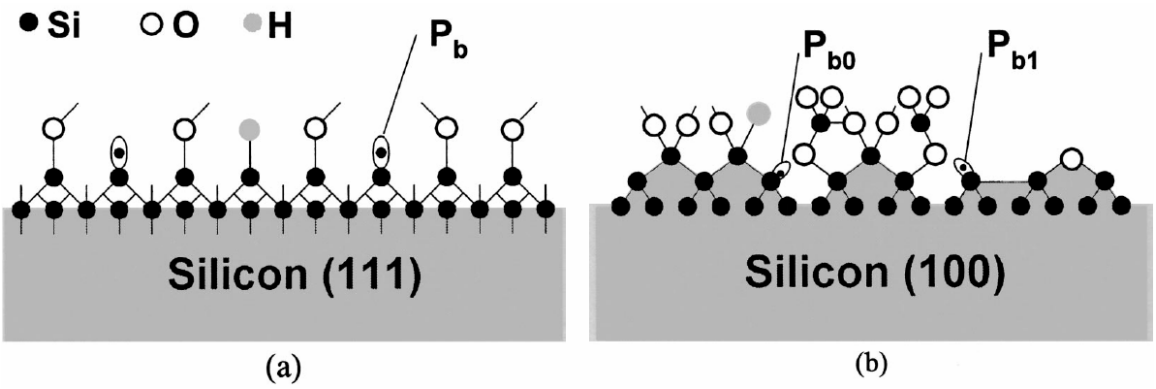


Figure 12. Interface States on the Silicon [17]**.**

Although not all of these P_b states are electrically active, those that are provide a trap that can capture holes or electrons. The effect of these creates a distribution of possible states that can be occupied in the exclusion regions of the band gap. In the upper part of the band gap, the traps act as acceptors so that they will be negative when filled and neutral when empty [15]. In the lower region of the band gap, the traps act as donors. As such, they will be neutral when filled and positive when empty [15]. The following figure shows the representation of the interface states in the gap and the charge contribution to the device.

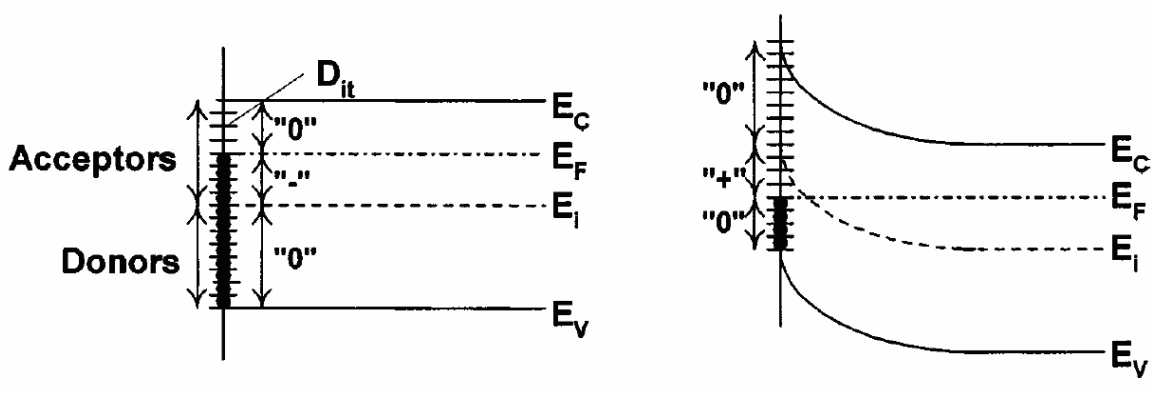


Figure 13. Interface States in the Band Diagram [17]**.**

The effect of the trapped charge is concentrated at the interface of the silicon and its oxide. This effect is similar to the results of the fixed charge and changes the voltage by the following amount [15, 17].

$$\Delta V_G = -\frac{Q_{IT}(\Phi_S)}{C_o} \tag{3.10}$$

The Φ_S term indicates that the charge is accumulated and measured at the surface of the channel.

These non-ideal effects are what lead to the generation of the NBTI. The factor that contributes the most is the generation of interface states due to process dependant factors.

3. Overall Effect on the Threshold Voltage

All of the non-ideal effects combine to form a more accurate model of the device. If the device has all of the discussed effects from above, the resulting change in the threshold voltage is presented below.

$$V_T = 2\Phi_F - \frac{K_S}{K_O} x_o \sqrt{\frac{4qN_D}{K_S \varepsilon_o}(-\Phi_F)} + \Phi_{MS} - \frac{Q_F}{C_o} - \frac{Q_{IT}(\Phi_S)}{C_o} \tag{3.11}$$

If the interface states and fixed charge were to increase in magnitude, the threshold voltage will also increase in magnitude.

C. ORIGINS OF THE NBTI DEFECT

As it was introduced above, the oxide that grows upon the silicon crystal leaves traps for holes or electrons. However, due to the technique that is used to grow the silicon dioxide, hydrogen is introduced into the interfacial areas and the device in general. The hydrogen bonds with the available dangling silicon bonds and effectively passivates the trap. Most experimentally observed data indicates that the instability arises as a result of a chemical reaction of the hydrogen at the interface and a subsequent diffusion of it through the oxide [11, 19]. As the hydrogen disassociates from the silicon bonds, the interface traps are exposed. The shift in the threshold voltage is dependant on both the stress time and temperature, but it specifically relates to the diffusion of the

hydrogen from the interfacial surface. It is also suggested that similar effects can be observed from the breaking of silicon and oxygen bonds in the oxide. Here the explanations for the NBTI will be discussed.

1. Reaction Diffusion Model

The representation that is used to describe NBTI is a simple reaction and diffusion model [11]. The reaction half of the model deals with the generation of interface states through the interaction of channel holes and the applied stress electric field. The other half relates to the transport of the hydrogen from the interface and its impact on the timing of the deterioration of the device performance. There are multiple methods to describe the reaction and diffusion model. Different approaches use different species of hydrogen, and there is still a decent amount of controversy over which is correct. However, neutral H_2 has a large amount of experimental data for support, and as such it will be the method presented.

a. Reactions

The generation of traps can occur in multiple ways. For the generation of the neutral H_2, a positively charge hydrogen atom will react with a hydrogen bonded to a silicon dangling bond. The reaction for this is shown in the following equation [17].

$$Si_3 \equiv SiH + H^+ \rightarrow Si_3 \equiv Si\bullet + H_2 \tag{3.12}$$

The tetrahedrally bonded silicon with the hydrogen termination is depassivated with the proton $\left(H^+\right)$. A source of protons can involve trapping holes from the channel with the following equation [17].

$$Si_3 \equiv SiH + h^+ \rightarrow Si_3 \equiv Si\bullet + H^+ \tag{3.13}$$

Another reaction can lead to the generation of neutral H_2. This involves creation of neutral H as it disassociates in the electric field, and is shown in the equation below.

$$Si_3 \equiv SiH \rightarrow Si_3 \equiv Si\bullet + H^0 \tag{3.14}$$

Other contributions can be added from fixed charge interactions, but the previous equations represent the majority of contributed reacted hydrogen to the reaction diffusion model.

b. Diffusion

If the main species for NBTI are neutral, their flux will be governed by diffusion. This occurs due to a difference in concentration and is related to the mobility if hydrogen through the different materials. As they depart from the interface, they leave behind a concentration of interface states. The flux of the hydrogen is defined in the following equation [11].

$$J_H = -D_H \frac{dN_H(x,t)}{dx} + \mu N_H(x,t) qE_{OX} \tag{3.15}$$

In Equation 3.15, the D_H is the diffusion coefficient of hydrogen through the oxide, N_H is the number of hydrogen as a function of position and time, μ is the mobility of hydrogen, q is the charge of the drifting species, and E_{OX} is the electric field in the oxide. The drift and diffusion components are further defined by the continuity equation as [11]:

$$\frac{dN_H}{dx} = D_H \frac{d^2 N_H}{dx^2} - \mu q E_{OX} \frac{dN_H}{dx} \tag{3.16}$$

The number of interface states (N_{it}) can be calculated from the total amount of hydrogen produced using the following relationship [11].

$$N_{it}(t) = \int_0^\infty N_H(x,t)\,dx \tag{3.17}$$

The current explanation to the diffusion assumes production of neutral hydrogen at the interface that combines to for molecular hydrogen. From this point the hydrogen concentration decreases in a linear fashion as distance increases from the interface. The following figure depicts this perspective.

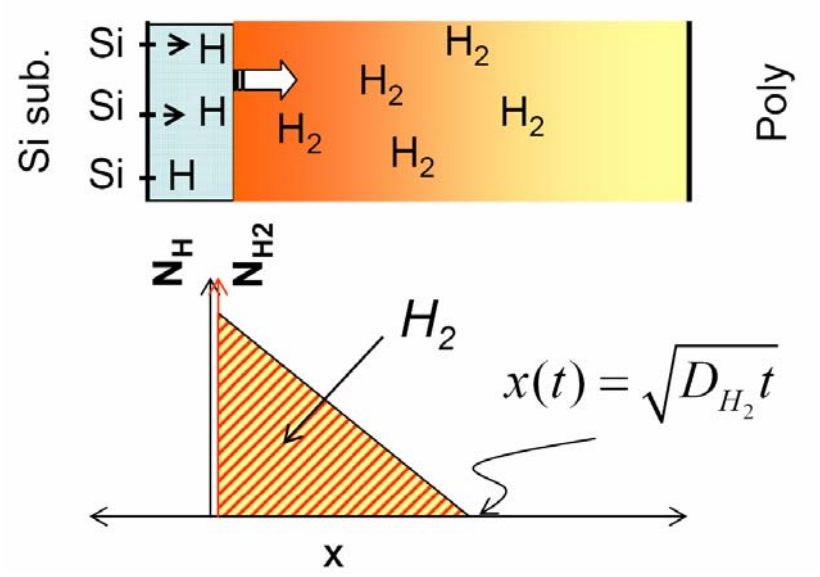


Figure 14. Diffusion of Hydrogen [19].

Due to the fact that these are both neutral species as mentioned before, the drift portion of Equation 3.15 is removed. The actual number of interface states is then calculated from Equation 3.16 using the diffusion coefficient of hydrogen, D_H. This yields the following [11, 19].

$$N_{it}(t) = \int_0^{\sqrt{D_{H_2}t}} N_H(x,t)\,dx \cong \frac{1}{2} N_{H_2}(0)\sqrt{D_{H_2}t} \tag{3.18}$$

Since the H_2 is formed of the available H, their concentrations are proportional such that $N_{H_2}(0) \propto N_H^2(0)$ [11, 19]. This leads to the final equation for the number of interface states shown below [11, 19].

$$N_{it}(t) \propto \left[\frac{k_f N_0}{2k_r}\right]^{\frac{2}{3}} \left(D_{H_2}t\right)^{\frac{1}{6}} \tag{3.19}$$

The valuable item of this equation is that of the exponent on the time value. This is what can be used to track the deterioration of the device as the stress time

increases. It follows the general form of the equation that relates the activation energy of the degradation, E_a, as well as the temporal component to the number of interface states. The form is shown in the following equation.

$$N_{it}(t) = Ae^{-\frac{E_a}{k_B T}} t^n \tag{3.20}$$

This equation, displayed in a logarithmic plot, shows the change in the interface states with time. This information is easily relatable to the change in the threshold voltage and to other circuit parameters. This also indicates that the devices degrade in time which can be magnified with higher temperatures.

c. Recovery

When the stress is removed, the mechanism for production of new hydrogen from the interface states stops as well. As a result, the device can recuperate a certain amount of its performance as hydrogen located close to the interface states will return to the dangling bonds from where they were removed. For this reason, circuits that operate under an AC stress will exhibit longer lifetimes.

2. PMOS Operation with NBTI Effect

Non-ideal operation introduces interface states between the bulk n-type channel and the insulating gate oxide as a stress is applied. According to Equation 3.18, the amount of interface states will increase according to a power law with an exponent of ~ 0.167. As the charge accumulates at the interface, a higher voltage in magnitude is necessary to maintain the same current. If an increased voltage is not applied, a device operating under an NBTI stress will slowly turn off as time increases. This is normally the condition that a component in a circuit will undergo, as the supply voltage remains constant over the life of a circuit. Such reductions can lead to timing errors of the circuit as well as inoperability in a worst case. The effect of the instability can be seen in the following figure of the drain current verses the drain voltage.

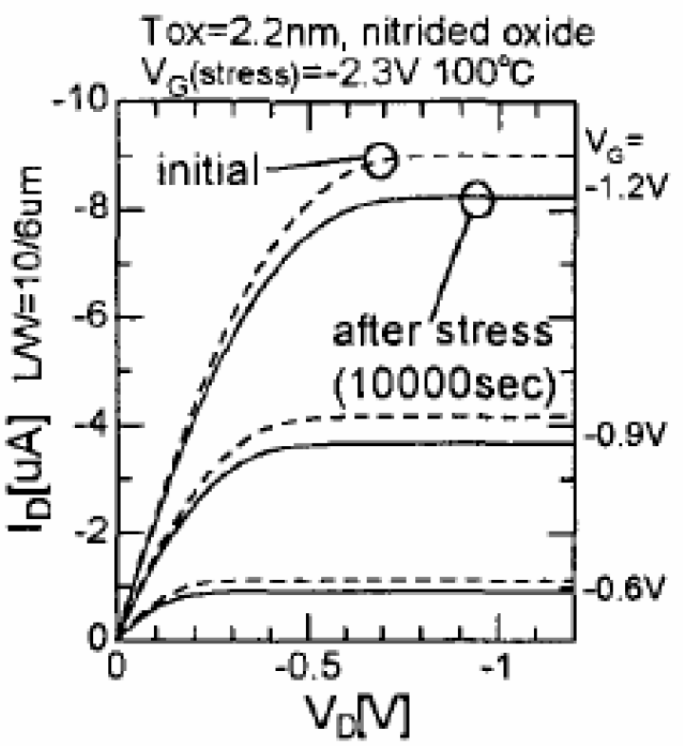


Figure 15. Impact of NBTI [19]**.**

IV. MEASUREMENT TECHNIQUES

This section presents the experimental techniques and procedures implemented to gather pertinent data. The first three sections introduce possible methods for extracting the device data for the NBTI. The fourth section relates to the experimental die and the temperature calibration of the thermistor data.

A. ON-THE-FLY MEASUREMENT

This technique is a widely used [1, 2] and is a simple method to evaluate the operational characteristic of PMOS devices. The technique biases the device such that it will be operating in the linear portion of the triode region at a stress temperature. This bias is held for the length of the test and the degradation of the device can be quantified as a percentage change in the device drain current I_{Dlin}. With this information, it is possible to determine the change in threshold voltage. The simplicity of this measurement makes it an attractive option, and this approach is implemented in these experiments. It is important to know about how the PMOS device operates to know how this test can be effective.

1. PMOS Parameters

Referring to the parameters discussed in the NBTI section on transistor operation, the On-the-Fly measurement technique biases the device into the linear portion of the triode region. Comparing the initial characteristics to the post-stress results reveals the departure from the original operation. The relationship of the voltages and the drain current are described in Equation 4.1 for the triode region [13].

$$i_D = k_p^{'} \frac{W}{L}\left[\left(V_{GS} - V_t\right)V_{DS} - \frac{1}{2}V_{DS}^2 \right] \tag{4.1}$$

In this equation, the terms are as follows. The width and lengths of the channel are represented as W and L. The $k_p^{'}$ is called the process transconductance parameter [13] that is related to the transconductance of the device. The rest of the values represent different voltages. V_{GS} is the gate to source voltage, V_t is the threshold voltage of the

device, and V_{DS} is the drain to source potential. If the drain bias is sufficiently small, the effect of the squared term $\frac{1}{2}V_{DS}^2$ can be removed from the equation. This reduces to the following linear current relationship [13].

$$i_D = k_p^{'} \frac{W}{L} V_{DS} \left(V_{GS} - V_t \right) \tag{4.2}$$

In this form, any change in the threshold voltage will generate a proportional change in the drain current. With some further algebraic manipulations, the change in the threshold voltage can be extracted from this relationship. The result is the equation presented by S. Mahapatra in his NBTI tutorial at the IEEE Reliability Symposium [1].

$$\Delta V_T = \frac{\Delta I_D}{I_{D_0}} \left(V_G - V_{T_0} \right) \tag{4.3}$$

In this form, the change in the threshold voltage is indicated by ΔV_T. The difference between original values of current I_{D_0}, and the degraded value are incorporated in the ΔI_D term. The final missing piece of information is that of the initial threshold voltage V_{T_0}. It can be determined by an extrapolation of the linear region of the V_{GS} versus I_{DS} curve to its 'x' intercept [14]. In the previous section, Figure 7 shows this approach.

2. Setting up the On-the-Fly Measurement

The main parameter for measurement is the linear drain current, I_{Dlin}. This is monitored as the device operates in the linear portion of the triode region. The experiment can be set up using Source/Monitor Units (SMUs) in the following fashion. The SMUs would properly bias the channel while stressing the gate and monitoring the drain current as depicted in Figure 16.

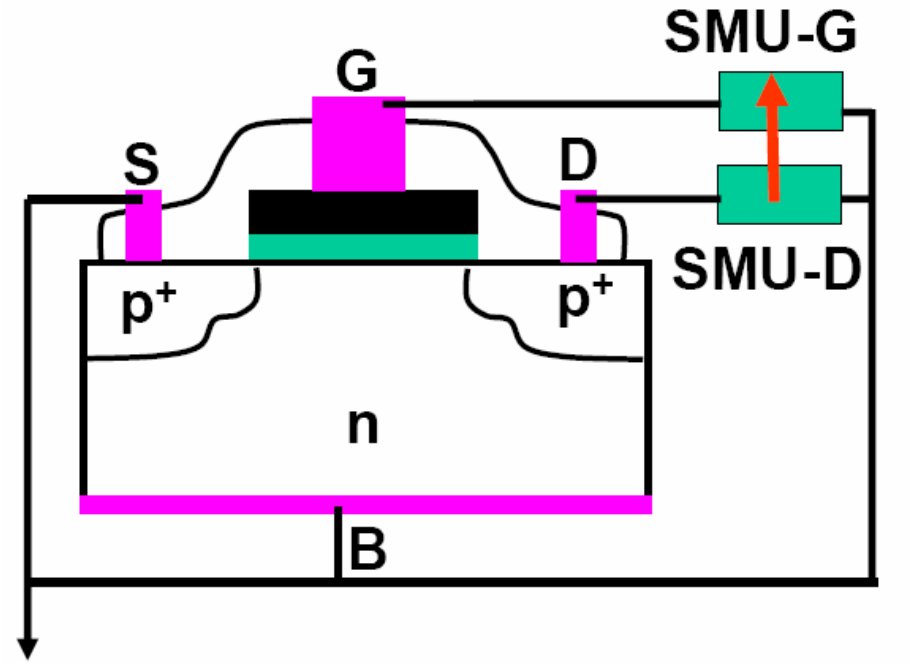


Figure 16. Example On-the-Fly Measurement Setup [1]**.**

The threshold voltage experiments described in the previous section provides guidance for extracting this data. Using Equation 4.3 from above, one determines the threshold voltage of the device. This is also assuming that the k_p' factor of Equation 4.1 does not change appreciably over the period of the stress.

3. Disadvantages of the On-the-Fly Measurement

Though this method provides a measure of the degradation of the device, it does not provide information as to the physical mechanism of the instability. Specifically, it does not provide the interface defect concentrations or any information about them [1]. It would be possible to use other techniques in conjunction with these to relate the interface data to the degradation, but as of yet there is no way to actively measure the interface data as the stress is present with this technique. Also, it is possible that the k_p' factor might change due to the fact that it is related to the mobility of the transistor as well as the capacitance of the oxide. It is possible that the charging might affect this value as well and make the device perform in a non-linear fashion.

B. CHARGE PUMPING

The charge pumping technique is a method for determining characteristics of non-ideal operation. Specifically, the technique can provide pertinent information about the

interface states at the Si-SiO_2 interface in the devices. It can provide data about the number of interface states, energy of the interface states, and the capture cross section [20]. This technique was attempted but was not used in this work due to difficulties encountered during the process. There are additional disadvantages listed below.

1. Technique and Theory

The source and drain of the test MOSFET are reversed biased as a stress waveform is applied to the gate. As the stress waveform transitions between the different states, the channel of the MOSFET will transition between accumulation and inversion. During this process, the interface states will be filled and vacated. Due to the bias on the drain and source, the charges will flow to these locations after the traps are vacated. They will be periodically filled as the transitions occur, being supplied with carriers from the connection at the bulk. This current that flows is the charge pumping current and is proportional to the number of interface states. The experiments all use the same experimental setup, but vary the stress waveform to gather different characteristics. Figure 17 shows the basic form of the experiment.

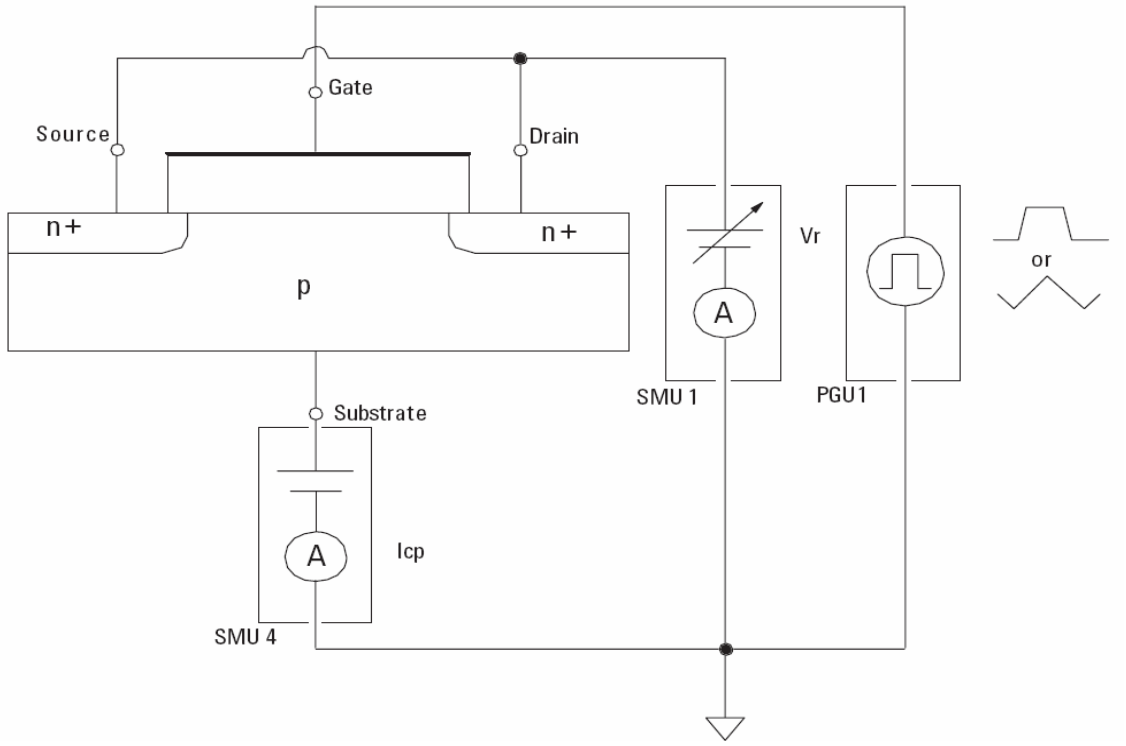


Figure 17. Charge Pumping Setup [20]**.**

2. Advantages

The different types of charge pumping all provide important data about the device. Specifically, the interface state information is a direct interpretation of the degradation of the devices. Likewise, this technique has been around and used since 1969 and is well understood [16, 21]. It is also possible that this experiment could be used to compliment the On-the-Fly technique by being run before and after the On-the-Fly measurement. In this fashion, it could provide interface concentrations before and after the other measurement.

3. Drawbacks

The major problem with using this technique is the delay associated with it. In order to test the device using this technique, the stress that is causing the NBTI must be removed. From the time that a stress is removed from the test device until the experiment is complete, the device relaxes some of the damage that has occurred. This temporal delay from going between a stress condition to a charge pumping test can provide erroneous results on device degradation [1, 2, 19]. This effect causes the intended DC voltage stress to appear as an AC stress due to application and relaxation. As the experiment is carried out and the time period increases, the interruptions become fewer and further apart and the data approaches the actual degradation characteristics due to logarithmic sampling [2, 19, 22].

C. DIRECT THRESHOLD VOLTAGE MEASUREMENT

Though this approach to measuring the threshold voltage, presented at the IRPS Conference [23], is quite novel, it is not implemented in this research. However, this would be an excellent area for verification of experimental data presented in this work. This technique interrupts the applied stress of the test in order to measure the threshold voltage. However, it does it in such as way so that the data is gathered close to a microsecond after the stress is interrupted.

1. Experimental Setup

This method uses an operational amplifier (OPAMP) as shown in Figure 18. The OPAMP is used in a fashion such that when the test is switched from the stress position to that of the measure position, it will attempt to zero the difference between the inverting

and non-inverting inputs of the device. To do this, it will force the necessary threshold voltage to the gate. The measurement delay that results then is only related to charging device capacitances [23].

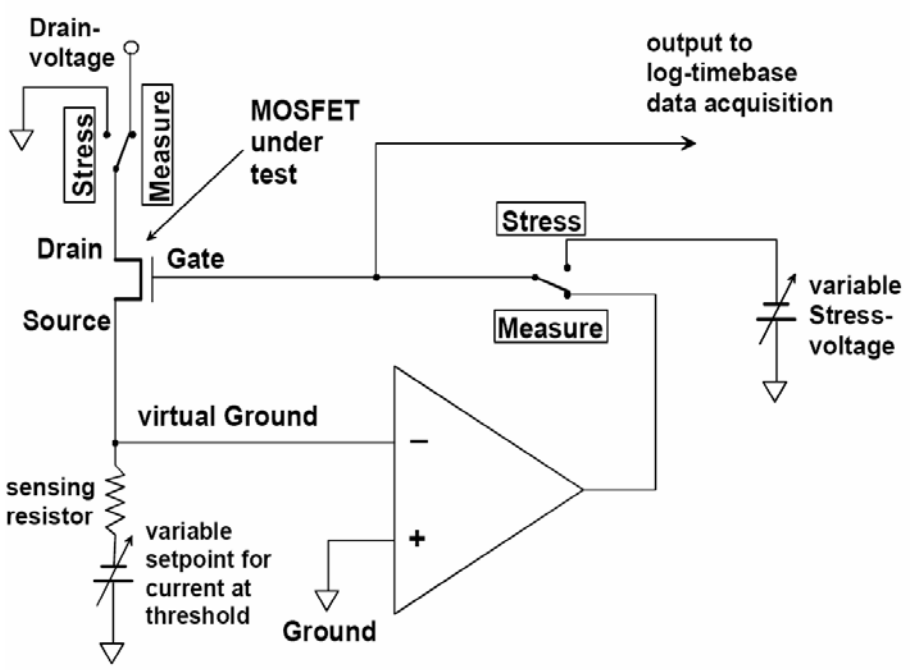


Figure 18. Ultra Fast V_T Measurement Setup [23].

2. Advantages

The most obvious advantage of this technique is the direct and immediate measurements made from the use of the OPAMP. As a result, it is possible for this test to be run on shorter time periods of stress. Furthermore, this approach allows a direct measurement of the threshold voltage.

D. THERMISTOR AND HEATER USAGE

All of the techniques mentioned above use an elevated temperature to accelerate the degradation. The usual tests require the use of a heater, such as a hot chuck on a probe station. For this thesis, the thermal input is generated by the integrated heater and thermistor combination. In order to obtain accurate measurements at the desired stress temperature, the behavior of the thermistors and heaters must be known. Ideally, the performance values of the thermistor and heater could be correlated to the p-n junction

diode current of the source or drain connection to the bulk for an accurate indication of the device's actual temperature. However, this is beyond the scope of this work and is included in other thesis research studies. This section discusses the technique used for calibrating the thermistor and heater to the thermal output of a Micromanipulator Heat Control Module and Hot Chuck.

1. Theory

The use of a thermistor and a heater exploits the approximately linear relationship of a material's resistance as the temperature changes. Though not all materials exhibit this property, certain metals do. With each material that fits this linear profile, there is a value related to the slope that can aid in monitoring the device's resistance and correlating temperature. This value is the Temperature Coefficient of Resistivity (TCR) with its relationship defined in the equation below [12].

$$\alpha_0 = \frac{1}{\rho_0}\left[\frac{\delta\rho}{\delta T}\right]_{T=T_0} \tag{4.4}$$

The ρ_0 term is the resistivity of the material at the initial reference temperature T_0. The partial derivative $\frac{\delta\rho}{\delta T}$ indicates a change of the resistivity from the initial ρ_0 to another ρ as the temperature changes from T_0 to T. This equation can be manipulated into the following form of Equation 4.5 with a linear approximation.

$$\rho = \rho_0\left[1+\alpha_0\left(T-T_0\right)\right] \tag{4.5}$$

For this to be useful, the relationship of resistivity and resistance shown below can be incorporated into the above equation to yield the other useful equation. This also assumes that the area and length of the resistor remain constant over the range.

$$R = R_0\left[1+\alpha_0\left(T-T_0\right)\right] \text{ where } \rho = \frac{RA}{L} \tag{4.6}$$

The TCR can be experimentally determined by measuring the resistance at known temperatures. For this test series, Equation 4.6 now allows the measured resistance of the thermistor to be correlated to the heater output of the hot chuck.

2. Procedure

The setup is displayed graphically in Figure 19. First, a bare die is placed upon the surface of the Micromanipulator Hot Chuck. Electrical connections are made to the Kelvin bond pads of the thermistor to determine the resistance by way of the Signatone probes to the HP4155B. A voltage is applied to the force connections, which are pads four and fourteen in Figure 19 and the differential voltage across the sense lines is measured. The resistance of the thermistor can be extracted from this method by dividing the resultant voltage read from the sense lines by the current applied to the force lines. There will be self heating due to the application of a current, but if the same amount of voltage is applied at each point of data extraction and the power is small in wattage in comparison to the applied heater wattage, the error can be negligible. It is also possible that the dissipated energy dropped across the resistance is not entirely converted to thermal energy, such as in an electro-migration instance. This procedure is carried out at different temperatures to yield a relationship of the thermistor resistance to the temperature.

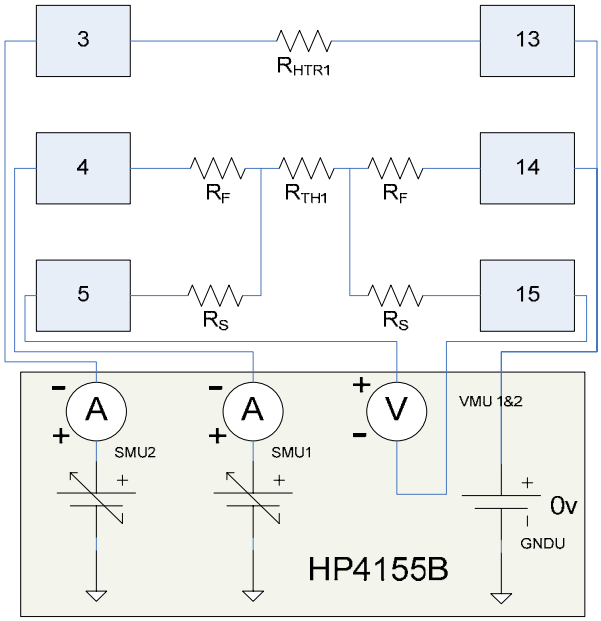


Figure 19. Heater Test Circuit.

The resulting data can be plotted and the linear relationship used to determine the TCR. The appropriate resistance can then be calculated to provide an accurate measurement of the device temperature. For different thermistors on other devices, it can be assumed that the TCR will be the same as they are made of the same material. Therefore, an initial reading can be taken at room temperature to determine the preliminary data point and the resistances at desired temperatures above room temperature can be extrapolated from this point with Equation 4.6.

THIS PAGE INTENTIONALLY LEFT BLANK

V. RESULTS AND CONCLUSIONS

The experimentation procedures and approaches are featured in this section. The applied techniques are highlighted and explained with the results of the experimentation. There are two distinct aspects to the configuration, that of the heater calibration to the thermistor, and that of the actual stress tests related to NBTI.

A. HEATER AND THERMISTOR RESULTS

A multi step process was used to extract the important data from the heater and thermistor combination. First, the TCR was determined from a selection of devices for the thermistor. Next, the device to be tested is measured at an initial condition and the temperature dependence is extrapolated using the TCR. Lastly, the heater is biased to reach the desired temperatures by relating the thermistor reading to the temperature. The results from the tests are provided.

1. Thermistor TCR

Four separate die were used to determine the TCR. The thermistor on the first die was measured at three temperatures to determine if the resistance increased in a linear fashion with temperature over the desired range of experimentation. It was tested at 20˚, 75˚ and 100˚C. Three other devices were measured at 25˚ and 100˚C for comparison. Figure 19 shows the manner in which this experiment was set up. The die is placed on the hot chuck and the resistances are sensed with the SMUs. The results are summarized in the following table.

Table 1. Resistance Measurements.

	Average Resistance (Ω) at Temperatures:				
Device	**20˚C**	**25˚C**	**75˚C**	**100˚C**	TCR
A	23.99	-	28.00	29.60	0.0029
B	-	23.84	-	29.34	0.0031
C	-	24.80	-	29.63	0.0026
D	-	24.55	-	29.60	0.0027
Average		24.40		29.54	0.0028

The data in Table 1 are presented as averages. As the thermistor was raised to the temperature listed via the hot chuck, the resistance was measured for a period of 315 seconds. The values reported in the table are an average of the resistance values measured, excluding the initial five seconds. The row entitled "Average" indicates the averages of the resistances for the different dies. The average value for the TCR will be used to relate the temperature dependence for the rest of the experiments. It is possible from this data to gather that the TCR could vary from plus to minus 10% as observed from the mean value and therefore the temperature may vary in that range. This can be seen by some calculations with the provided data. Using Equation 4.6 and solving it for the temperature yields the following equation.

$$T = T_0 + \left(\frac{R}{R_0} - 1 \right) \frac{1}{\alpha_0} \tag{5.1}$$

Using the data from Table 1, it can be shown how the 10% change in the TCR will affect the temperature reading. For this explanation, it is assumed that the TCR is the only component that changes.

$$\Delta T = \frac{T_0 - T_M}{T_M} \times 100\% \tag{5.2}$$

The T_M value is the actual measured temperature and the T_0 value is the temperature using the mean value of the TCR measurements. For the example of Device B, it is known that the device is at 100C. Calculating the temperature with the mean TCR yields a temperature of $T_0 = 107°$. This gives an error of $\Delta T = 7.4\%$ indicating the mean value predicts a higher temperature than the actual. Using Device C, the error is $\Delta T = -5.4\%$. The implication of this is that if the TCR is selected to be the mean value, the actual temperature will differ from the actual value by a value on the order of 10% above and below that temperature.

2. Heater Bias Determination

The heater operation is characterized using the same experimental setup as in Figure 19. Typical biases used voltages below two volts to locally heat the devices. Typical currents would be less than $0.1A$ as the resistance of the heater was near 20Ω.

The heaters were characterized for one wire bound structure and correlated to the thermistors on the device. This device provided the experimental stress data as well. To begin the process, the resistance of the thermistor was measured at room temperature. Next, the TCR is used to extrapolate the temperature dependence. Then the heater voltage is stepped from 0.0v to 1.75V and the resistance of the thermistor is recorded. The results are displayed in the table below.

Table 2. Heater Bias and Resistance Measurements.

	Device A		Device B	
V_{HTR} (V)	R (Ω)	T (˚C)	R (Ω)	T (˚C)
0.000	21.995	20.100	21.713	20.300
0.250	22.124	22.165	21.773	21.276
0.500	22.504	28.264	22.143	27.285
0.750	23.109	37.964	22.779	37.628
1.000	23.916	50.914	23.595	50.882
1.250	24.883	66.424	24.570	66.734
1.500	25.978	83.995	25.683	84.821
1.750	27.188	103.416	26.898	104.559

After these initial characterizations are completed, the biases to achieve the desired test temperatures are calculated. The following table shows the resultant biases.

Table 3. Heater Bias Results.

	Device A			Device B		
T (˚C)	P_{HTR}	V_{HTR}	R (Ω)	P_{HTR}	V_{HTR}	R (Ω)
25	0.009	0.381	22.301	0.008	0.363	22.002
75	0.089	1.418	25.417	0.083	1.369	25.079
100	0.129	1.739	26.975	0.121	1.680	26.617

These biases can now be used to heat the device to the target temperature within a reasonable degree of accuracy. Thought this relies upon the linear effect of the TCR, it is possible that these calculations are not entirely accurate. However, this should place the device at a temperature close to the desired value.

B. STRESS MEASUREMENT RESULTS

Following the heater and thermistor tests, the devices are ready to undergo the $I_{D_{lin}}$ degradation test. The first step in the process is to determine the threshold voltage from the V_{GS} versus I_{DS} curve. After this is complete, the stress test can be performed at the different temperatures.

1. Threshold Voltage

The following plot displays the connections made to extract the threshold voltage from the device. The SMUs were connected to each of the four contacts as the test is run. This shows the connections for the first structure, and is the same for the second device that has a built in protection diode.

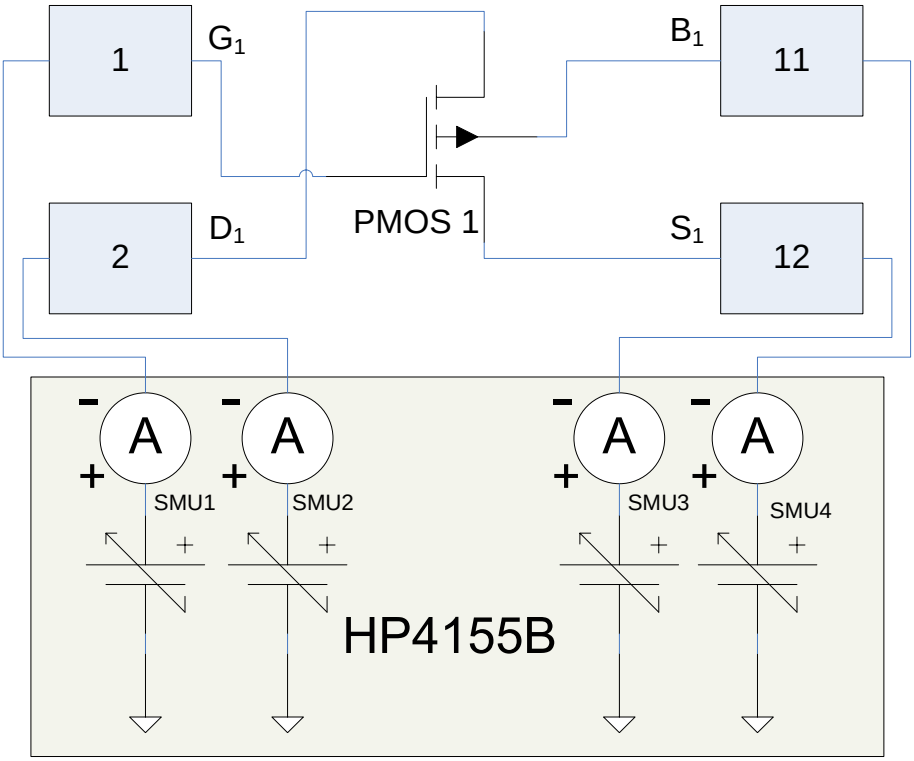


Figure 20. Threshold Voltage Test Experimental Setup.

The following graphs depict the current and gate voltage plots produced from the devices. The resultant values are adequately close to the threshold listed from the IBM data. The first device has a threshold of $-0.271V$ while the other has a threshold

of -0.298*V* . They differ by 0.027*V* . These are listed as the intercepts of the line that is drawn tangentially to the linear region on the figures. These data points will serve as the starting point for the devices and degradation will be metered from this point.

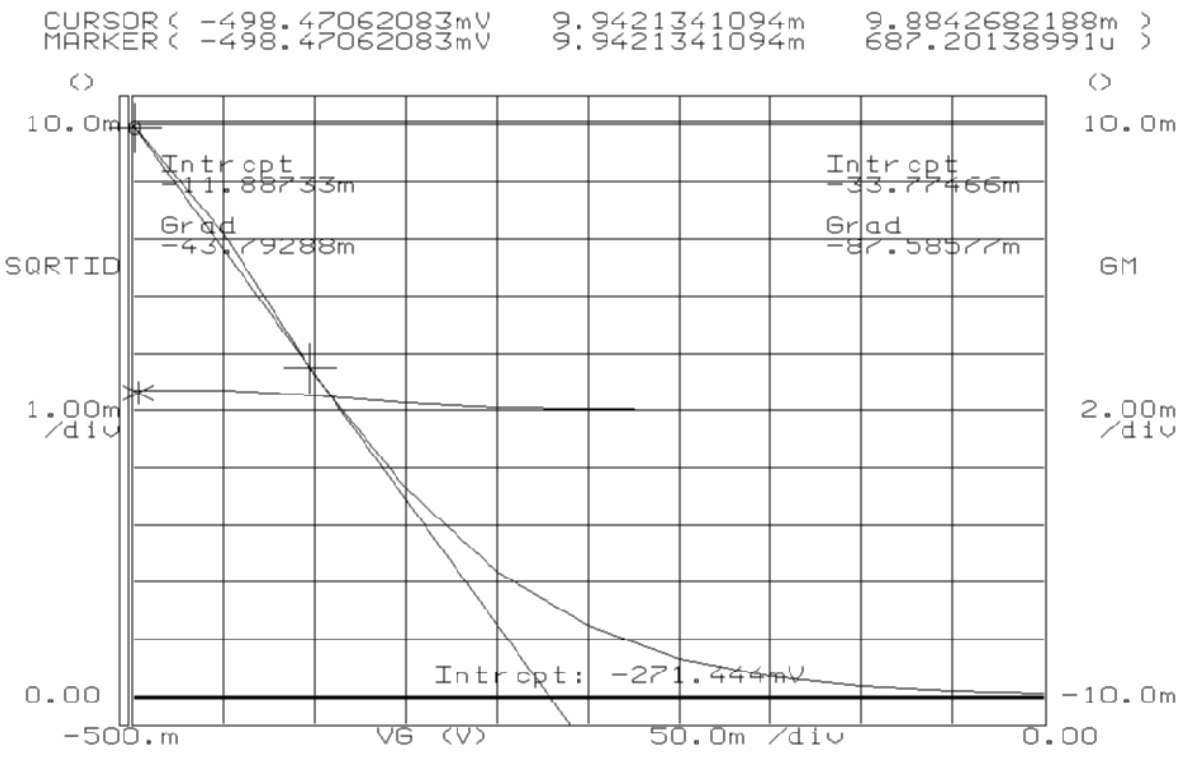


Figure 21. First Device Threshold Voltage Measurement.

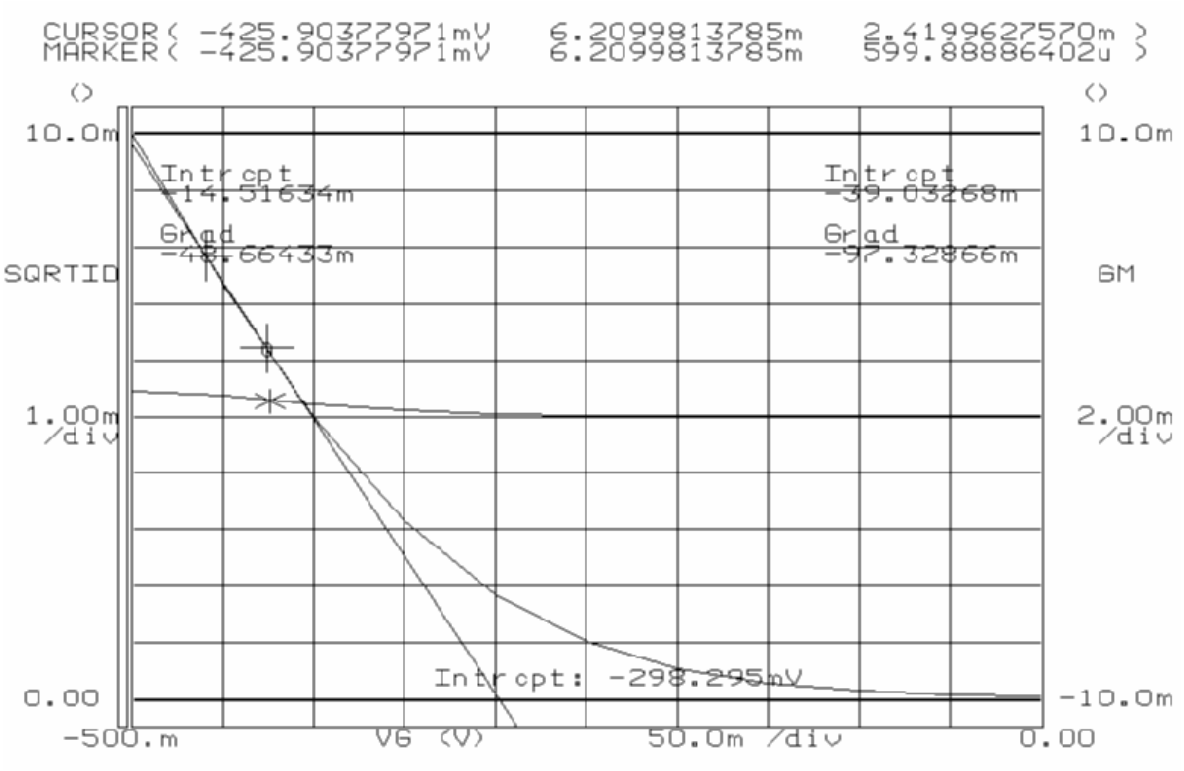


Figure 22. Second Device Threshold Measurement.

2. Stress Experiment

The experiment is set up according to Figure 23. The SMUs are set to force the desired voltages for the stress as well as bias the heater and thermistor for temperature control. While the units force the voltages they also monitor the currents for the analysis.

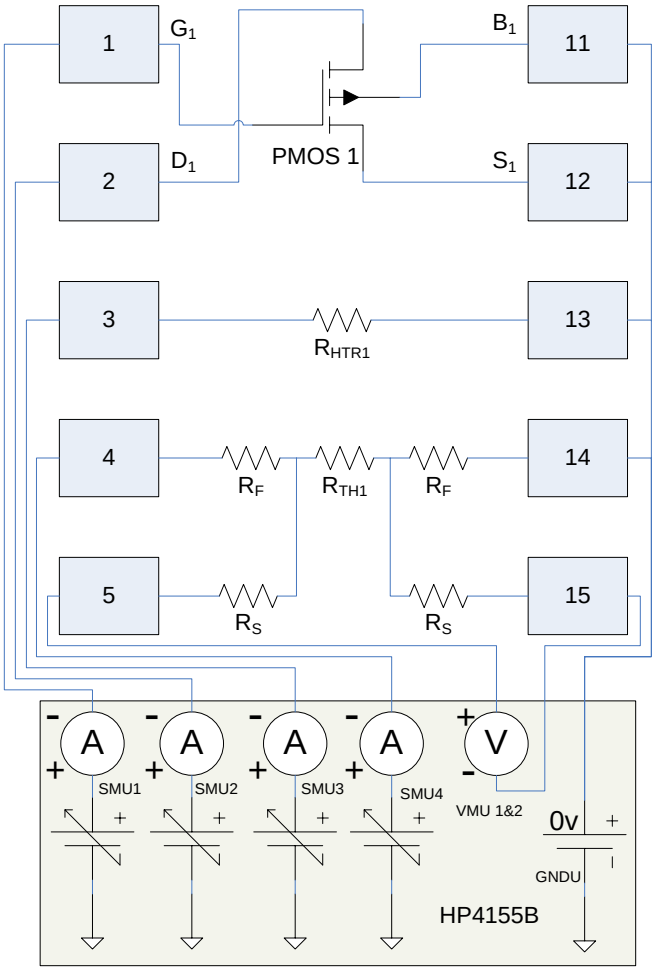


Figure 23. Test Setup.

The tests themselves were conducted separately by way of using the software available for the HP4155B. The stress values were chosen to be a gate stress voltage of $V_G = -2.2V$ and a drain voltage of $V_D = -0.2V$. The resulting value of drain current in this state was in near $I_D = -1.6mA$. The experiments ran at the two chosen temperatures of 25˚ and 100˚C. Initial tests at 25˚C applied the stress for a period of one hour. The tests ran at 100˚C were taken at two different intervals. The first ran for a period near three hours, and the final test ran for near 8 hours. Both devices were subjected to the stress at 25˚C prior to the stress at 100˚C. The results are shown in the following plot that show the change in the linear drain current as the device operates. In the legend, the first number indicates the temperature of the experiment. The second number indicates the die number, which in this case is die number three. Die number three is so named due

to the fact that it was the third wire-bound device to undergo experimentation, but it was the first to be tested with the On-the-Fly measurement. The third number specifies which transistor on the die was used, and the final number identifies the first or second run of the experiment. The later values are calculated using the initial value of the drain current as the initial condition.

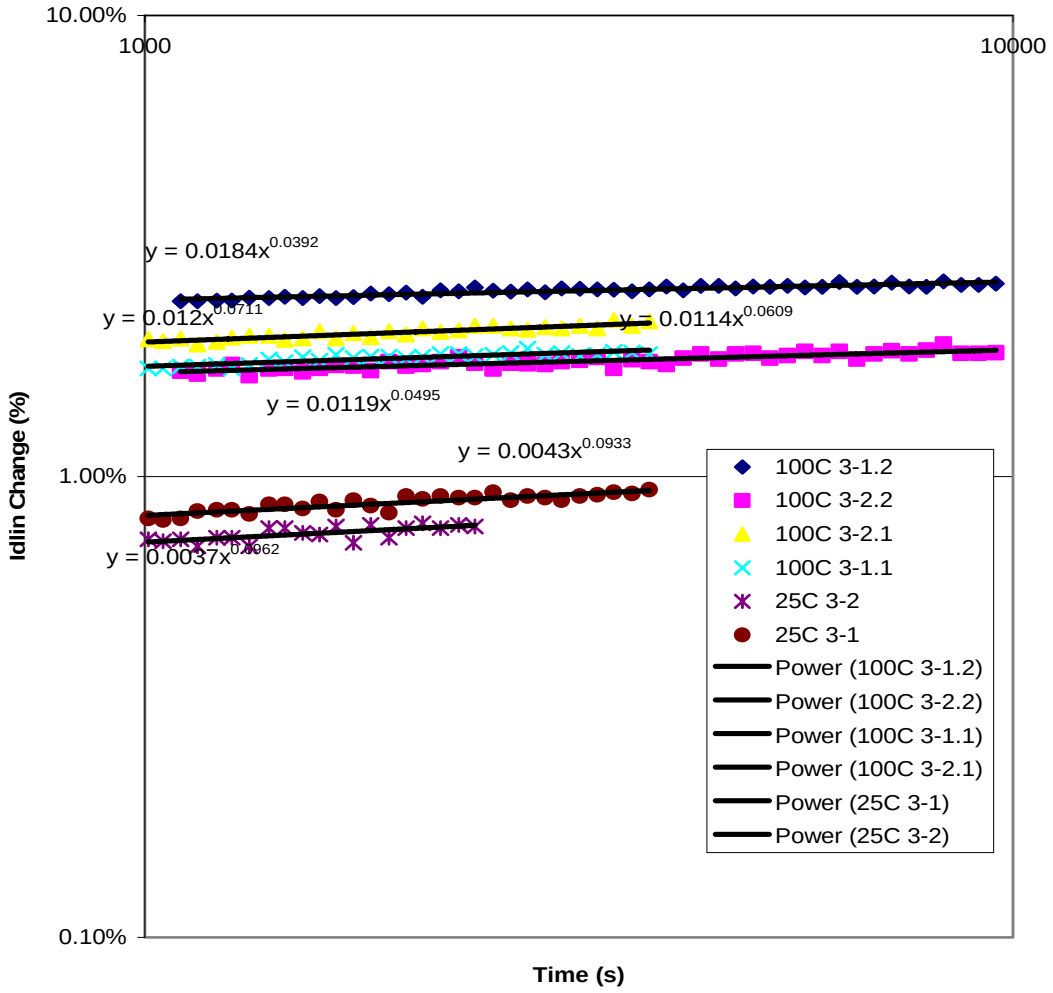


Figure 24. Stress Results.

The results have a power law fit to the data, and the exponent should be an indication of the specific type of the hydrogen diffusion. These values are then used to generate the resultant shift in the threshold voltage. These plots are generated using Equation 4.3 from the previous chapter. Due to the linear relationship between the drain current and the threshold voltage, the exponents will be the same for each portion of data.

The values only act to scale the data to the threshold voltage values. The following graph depicts the threshold voltage change as the experiment time passes.

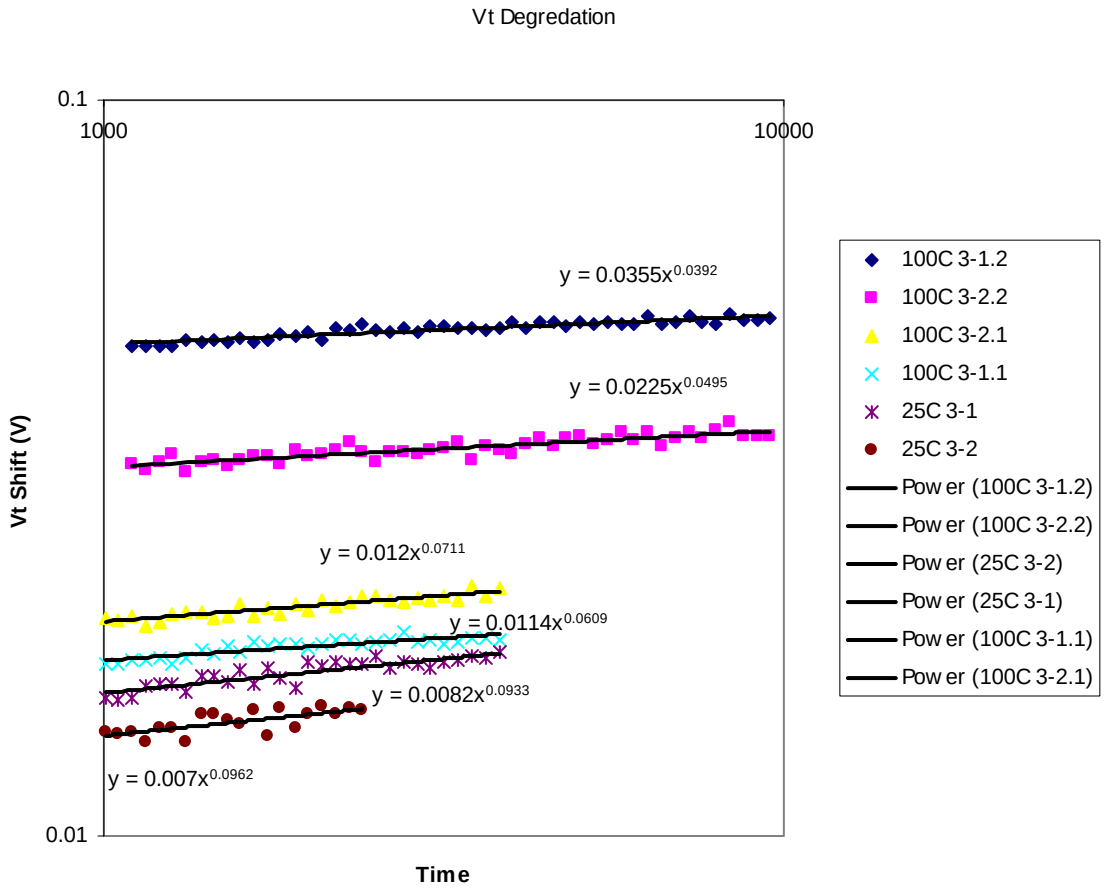


Figure 25. Threshold Voltage Shift.

It is also important to note that the graph only includes the time between the periods of 1000 to 10000 seconds. These were the points between which the heater current was stable and the temperature could be verified to be within the desired range. This is depicted below in the following graph. It shows that initially the temperature needed approximately 1000 seconds to stabilize. It also shows that as the heater was run at higher temperatures for an extended period of time, the performance degraded. If this is incorporated into the drain current plots, it would add an additional saturation characteristic [1].

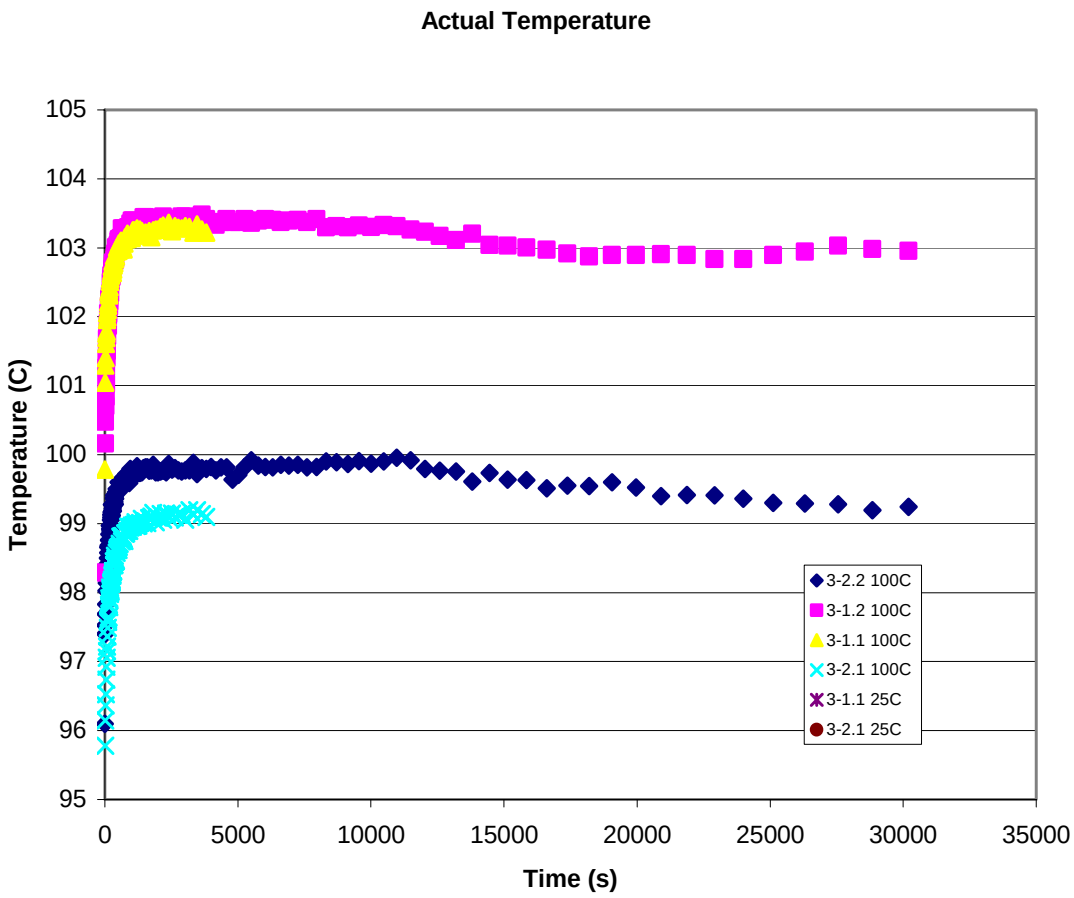


Figure 26. Temperature Results.

After the tests were completed, the threshold voltage measurements were run again to determine if there was an appreciable change. These were taken within the hour after the stress, and there may have been more relaxation past this point. These measurements are more of a proof of concept that there was a permanent or semi-permanent shift in the values. Some literature presents recovery data to determine time needed for the threshold voltage to settle to a final position [24]. These post stress results are exhibited in the following figures in the same fashion as Figures 19 and 20. The intercept of the tangential lines with the x-axis represents the threshold voltage. It is apparent that the devices did experience a shift in the threshold voltage. The first device shifted by $3.6mV$. The second device shifted by $10.9mV$. Both shifted in the proper direction as well, making the threshold voltage more negative and in effect, turning the device "off".

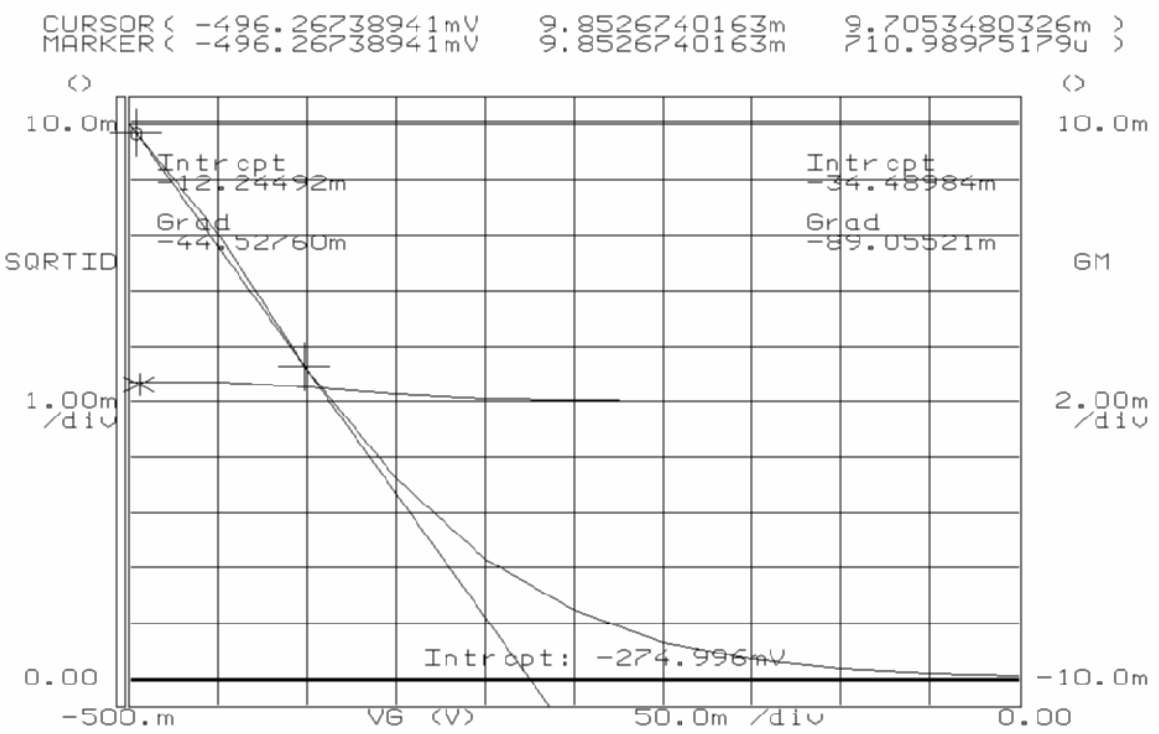


Figure 27. First Device Threshold Voltage Post-Stress Measurement.

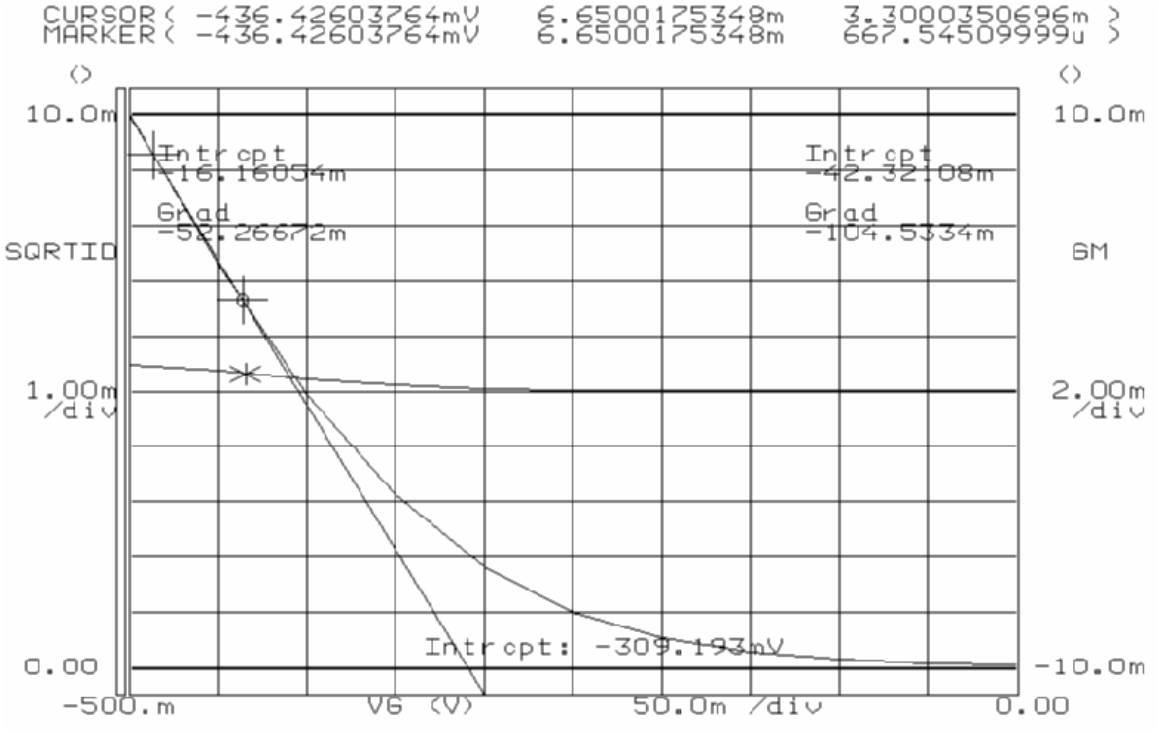


Figure 28. Second Device Threshold Voltage Post-Stress Measurement.

Now that the data is gathered it will be analyzed in the sections below. It is necessary to compare the experimental data of this test to the generated data in current publications to determine its validity.

C. INTERPRETATION OF RESULTANT INFORMATION

1. Heater Setup

There are a number of possible problems with the heaters as they were used in this experiment, but these areas can be improved in subsequent tests. For instance, the temperature of the heaters needs to be verified for a more accurate assessment of the devices' susceptibility to NBTI. The temperature could be correlated to diode current at a certain temperature to get a device temperature at the transistor level. Another problem area is of the heater control. In tests of longer duration, the performance of the heaters degraded. If this process was automated into a control loop, the temperature could be more accurately maintained.

The reasons for this saturation of the heater and subsequent reduction in temperature as seen in the previous figure could relate to the heat capacity of the mounted device. Eventually as the device heats the entire structure as well as the 28 pin DIP that carries it, the structure will reach a thermal equilibrium. It is possible that at this point there is some amount of leakage current that begins to bypass the heating element and cause the reduction in temperature.

2. NBTI Data

The data that is presented in this work does not match with the classical results that current published papers present. Classical data exhibits a power law dependence that yields an exponent in the range of 0.16-0.25 [1, 2, 19]. The data presented here ranges from 0.04-0.1. This is overly optimistic in nature and can indicate that this data is erroneous. At these rates, the NBTI worst case degradation only changes the threshold voltage by 10% in 68 years. They did show a change in the threshold voltage, as is seen in figures and in the results section. For one case, the threshold voltage shifted by near $0.01V$, and in the other case, it shifted by $0.004V$. Ideally, the results would be more consistent with published data. The following graph shows current published experimentally determined results for an NBTI stress.

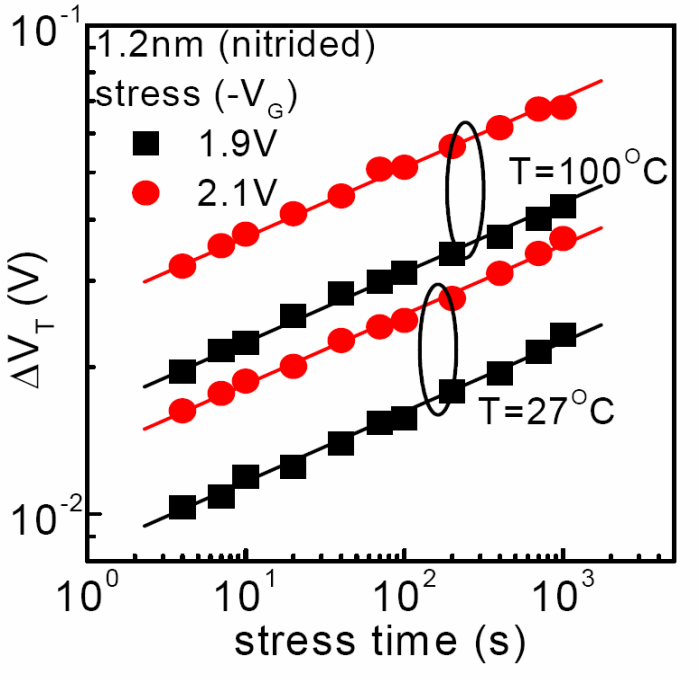


Figure 29. Classic NBTI Results [1].

It is not entirely clear why these experimental results came out as they did. There was a control issue with the heater that contributed to the data being off target. There was some amount of saturation of the devices, in that the damage decreased as the stress time increased. This can be attributed in part to the heater control issue. The exponents were relatively within agreement between the 25˚C measurements and that of the 100˚C case. More experimentation is necessary to determine the cause of the erroneous data. These and other reasons for problems are discussed in further detail.

D. CONCLUSIONS

Due to the variation of the experimental data relative to the published work, it is possible that there was some error in the setup and testing of the devices. The results are mildly perplexing in that they do show deterioration in the operational characteristics, but it is in such a reduced fashion that it would not affect device performance in any short period of time.

1. NBTI Test

In most experimental cases, the perpetrator of experimental error is the time delay in the experiments [1, 2, 22, 23]. This is a result of the interruption of the stress to test the device. This period of interruption allows the device to recover and usually gives an

optimistic exponent for the data. The On-the-Fly Technique is designed to eliminate the time in between the samples and give a better representation of the effect. However, this data does not show the promising features normally observed with this.

a. Thermal Issue

A portion of the erroneous data stems from an error in the temperature control, but this is not the sole source of error. As the heater drifted during normal operation, the exponent of the fitted power law line would decrease as a result of apparent saturation. This was observed in the data, and this agrees with published experimental phenomena. This difference in the power law slopes due to different temperatures is observed in other published experimental data as is shown below in the following graph. They depict an increase in slope with an increase in temperature, so, therefore, a reduction in temperature during operation would appear as a reduction in the slope.

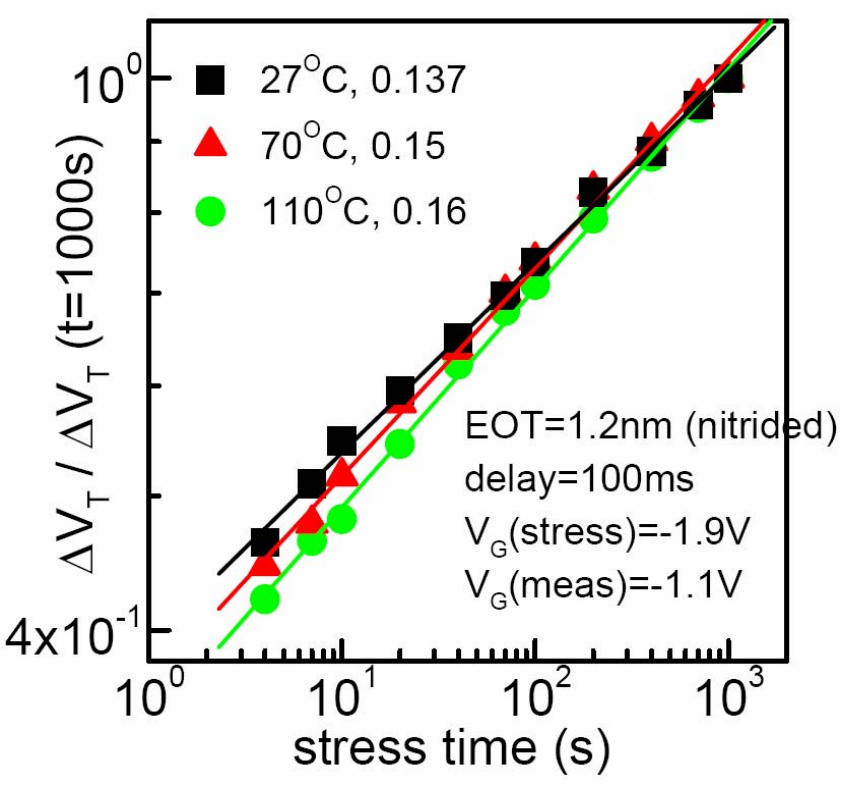


Figure 30. Effect of Different Temperatures on NBTI Results [1].

Though this did explain some problems experienced during the process, it does not entirely explain why the slopes are as low as they are. Furthermore, for the

observed data, the slopes of the higher temperature should be greater than those at the lower temperature. This is not the case for this presented data. The exponents for the lower temperature appear as greater than for the higher temperature case. Therefore, more problem areas must be considered for an explanation.

b. Threshold Measurement

There are many techniques for measuring the threshold voltage of a device. The method presented here may not be as consistent and repeatable as other methods, and may differ from IBM's process for threshold measurement. For instance, the values for the threshold voltage that were measured in this work are not too consistent even though they are found on the same die and were manufactured at the same time. This would introduce error in the values but would not necessarily impact the slope of the lines.

c. Drain Voltage Selection

Another possible problem area could be with the selection of the drain voltage during the On-the-Fly measurement. The value chosen for this experiment was $0.2V$, which is almost to the level of the threshold voltage. This could mean that the operational point was not that of the linear portion of the triode region. As a result, the necessary equation for calculating the threshold voltage shift would change. Some published sources indicate that appropriate voltages are more in the range of $0.05V$ or $0.1V$ [1, 2].

d. Additional Problem Areas

The problem could be one of erroneous connections or bad contacts or processing of the die. During the wire binding process, the devices are heated to 160°C. It is possible that a good portion of the available hydrogen that was passivating the interface states diffused from those states. This seems unlikely due to the fact that there was no field applied to these devices during the process. Though the data seemed erroneous, it still; did exhibit a reduction in performance that is usually associated with NBTI. Since there was a measurable shift in the threshold voltage, this must mean that there was a change in the number of interface states, N_{it}. This value of this change, however, was

not directly measured. These devices can be considered to show NBTI, and with further experimentation, will most likely exhibit the classic exponents published in present literature.

2. Areas for Further Study

The experiment needs to be rerun to determine in a more accurate fashion to what extent the devices are susceptible to NBTI. This initial work does give positive indication that there is some degradation, but more work is necessary. There are a couple of ways in which this should be achieved.

a. Heater Control

First, the best course of action is to establish a control algorithm or even a hardware device to control the temperature of the device. This will affirm that the temperature is not drifting from the desired value. Once this is assured, then there will be less areas of possible error. Initially this was to be completed by another student prior to this research, but it was not finished in time for this research.

b. Experimental Approach

The devices can be tested using any of the other methods for stress evaluation, such as the charge pumping method or the direct threshold measurement technique. This can be used for additional verification of the results presented here. One beneficial area for investigation in this vein would be to run the charge pump measurement to determine an initial N_{it}. After this, the $I_{D_{lin}}$ On-the-Fly measurement could be run for a length of time on the order of 6-8 hours. At the conclusion of the test, the charge pump measurement could be run again directly following the experiment and at different time intervals from that point. By running the experiment in this fashion, the active deterioration can be measured in conjunction with a verification of a change in N_{it}. Furthermore, running the charge pump after the stress is relaxed can give an indication of recovery characteristics. If these are run and compared to the data, they can also be used to determine what portion failed during the $I_{D_{lin}}$ measurements.

In order to run the charge pumping experiment, more device information is necessary. The charge pumping experiment needs information on the physical

characteristics of the device, such as effective widths and lengths of the channel. One reason that the charge pumping technique was not used is because the applied voltages must be carefully selected in order to not damage the devices. Designs for tests must take into account the maximum operating voltages that the devices can handle.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

[1] S. Mahapatra, "Electrical Characterization and Modeling of Negative Bias Temperature Instability in p-MOSFET Devices," pp. 192, 2006.

[2] T. Nigam, "Lifetime Enhancement under High Frequency NBTI Measured on Ring Oscillators," in *IEEE International Reliability Physics Symposium*, 2006, pp. 289.

[3] W. Abadeer and W. Ellis, "Behavior of NBTI under AC Dynamic Circuit Conditions," in 2003, pp. 17-22.

[4] Boeing. (2002, APR 2002). B-52 Stratofortress Background Info. *2006* pp. 1.

[5] Department of Defense. (1996, 31 DEC 1996). Test method standard: Microcircuits.

[6] Defense Science Board Task Force. (2005, February 2005). High Performance Microchip Supply.

[7] H. Livingston, "The current state of the semiconductor industry and DoD weapon system dependence on off-shore products," in *Diminishing Manufacturing Sources and Material Shortages 2005*, 2005,

[8] Institute for Defense Analyses, "Resource & Support Analyses," vol. 2006, 2006.

[9] G. Carlson, "Trusted foundry: The path to advanced SiGe technology," in 2005, pp. 9-12.

[10] IBM Systems and Technology Group. (2005), 130-nm Foundry Technology Platform: CMOS, RF CMOS and SiGe BiCMOS.

[11] D. Pierce. (2005), Module description.

[12] S. O. Kasap, *Principles of Electronic Materials and Devices*, 2nd ed. New York: McGraw-Hill, 2002, pp. 745.

[13] A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 5th ed. New York: Oxford Universty Press, 2004, pp. 1283.

[14] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*, 2nd ed. New York: McGraw-Hill, 2003, pp. 644.

[15] R. F. Pierret, *Semiconductor Device Fundamentals*, 1st ed. Redding: Addison Wesley, 1996, pp. 792.

[16] D. K. Schroder, *Semiconductor Material and Device Characterization*, 2nd ed. New York: Wiley, 1998, pp. 760.

[17] D. K. Schroder and J. A. Babcock, "Negative Bias Temperature Instability: Road to Cross in Deep Submicron Silicon Semiconductor Manufacturing," *J. Appl. Phys.*, vol. 94, pp. 1-18, Jul. 2003.

[18] J. P. Campbell, P. M. Lenahan, A. T. Krishnan and S. Krishnan, "Atomic scale defects involved in NBTI [MOSFET reliability]," in 2004, pp. 118-120.

[19] M. A. Alam, "A Simple View Of A Complex Phenomena," in *Reliability Physics Tutorial Notes,* 2006, pp. 191.

[20] Agilent Technologies. (2000), Evaluation Of The Surface State Using Charge Pumping Methods. pp. 1.

[21] J. S. Brugler and P. G. A. Jespers, "Charge Pumping in MOS Devices," *IEEE Transactions on Electron Devices,* vol. 16, pp. 297, 1969.

[22] A. T. Krishnan, "Modeling NBTI: Kinetics to Circuit," in 2004, pp. 198-198.

[23] H. Reisinger, O. Blank, W. Heinrigs, A. Mulhoff, W. Gustin and C. Schlunder, "Analysis of NBTI Degradation- and Recovery- Behavior Based on Ultra Fast VT-Measurements," *IEEE IRPS,* pp. 448, 2006.

[24] J. P. Campbell, P. M. Lenahan, A. T. Krishnan and S. Krishnan, "NBTI: An Atomic Scale Defect," *IEEE IRPS,* pp. 442, 2006.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Todd R. Weatherford
   Naval Postgraduate School
   Monterey, California

4. Andrew A. Parker
   Naval Postgraduate School
   Monterey, California

5. David Alexander
   Kirtland AFB
   Albuquerque, New Mexico